CESifo WORKING PAPERS

6739
2017

November 2017

Unraveling Over Time

Sandro Ambuehl, Vivienne Groves

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6739*
Category 4: Labour Markets

Unraveling Over Time

Abstract

Unraveling, the excessively early matching of future workers to employers, is a pervasive phenomenon in entry-level labor markets that leads to hiring decisions based on severely incomplete information. We provide a model of unraveling in one-to-one matching markets for prestigious positions. Its distinguishing feature is that the market operates over an extended time period during which information about potential matches arrives gradually. We find that unraveling causes potentially thick markets to spread thinly over a long time period. In equilibrium, an employers desirability is correlated neither with the time at which they hire, nor with the expected productivity of their matched worker. Unraveling thus significantly redistributes welfare among employers compared to a pairwise stable match. We study policies that manipulate the availability of information about students and show that they are effective only if they provide a sudden surge in information. Our main application is the market for U.S. federal appellate court clerks, a significant input into the efficiency of the justice system. Consistent with the model, hiring times in our dataset are spread over a period of six months and are uncorrelated with the desirability of a judge as an employer.

*Sandro Ambuehl**
University of Toronto
Department of Management UTSC and
Rotman School of Management
105 St. George St.
Canada – Toronto, ON, M5S 3E6
sandro.ambuehl@utoronto.ca

Vivienne Groves
Stanford University
Graduate School of Business
Knight Management Center
655 Knight Way
USA – Stanford, CA 94305
vivienne.groves@gmail.com

*corresponding author

October 21, 2017

1 Introduction

Entry-level labor markets frequently suffer from the excessively early matching of workers to employers, a phenomenon known as *unraveling* (Roth and Xing, 1994). Faced with a cohort of prospective employees all graduating at similar points in time, some employers attempt to eschew competition by hiring promising students early. By doing so, these employers diminish the pool of applicants available for hire after graduation, prompting others to hire promising students early, too. The resulting mutual undercutting frequently leads to frenzied and chaotic markets, and to hires made up to multiple years before employment starts, to the dismay of many market participants.

An important example is the market for federal judicial law clerks, our main application in this paper.[1] Judges depend heavily on their law clerks (Avery et al., 2001); their role includes everything from the management of disputes in district courts to the drafting of opinions at the U.S. Supreme Court (Bonica et al., 2017).[2] Hence, an efficient selection and matching of clerks to judges is a significant input into the efficiency of the federal justice system.

In this paper, we develop a formal model of unraveling in a two-sided matching market. We have two objectives. First, we aim to illuminate the mechanisms and determinants of unraveling. The market we study takes place over an extended time period during which information about potential employees evolves gradually. For employers, equilibrium behavior thus involves balancing the benefits from waiting for better information before making a hiring decision against the increased competition they would face for remaining in the market at later points in time. We show that equilibrium hiring times are widely dispersed across time, that unraveling acts as a powerful redistributive device among employers, and that a potential employer's attractiveness is uncorrelated with the point in time at which they hire. Our model thus contrasts with the previous literature on unraveling, in terms of both assumptions and results. That literature almost exclusively considers models with

[1] Other examples of unraveling in entry-level labor markets include clinical psychology internships, dental residencies, optometry residencies and Japanese university graduates (Roth and Xing, 1994). More recently, unraveling has also been observed in the market for entry-level private equity bankers (Alden, 2014), for technology jobs (Ante, 2012), for faculty in philosophy departments (Kueppers, 2016), for postdoctoral positions in theoretical physics (Yu, 2017), for college athletes (Popper, 2014), for professional soccer players (Hautmann, 2013), and in several medical subspecialties such as the epilepsy and clinical neurophysiology fellowships (Vidaurre and Campbell, 2017).

[2] Also judges' decisions on matters such as whether to file a separate opinion or to dissent in a case are based upon the support they anticipate from their clerks. Moreover, law clerks are essentially the only persons a judge can talk to in depth about a case (Wald, 1990).

two periods, an assumption that severely restricts the opportunities for avoiding competition. For a subset of the literature's results, our model helps clarify which are robust, and which are artifacts of the two-period restriction. Additionally, we show how the apparent stark distinctness of models of unraveling (Halaburda, 2010; Du and Livne, 2013) can be understood as extremes on a continuum.

Second, we hone policy tools to address unraveling. We focus on policies that vary the way in which information about students is divulged into to the market. Our main insight is that such policies have an effect on the resulting match only if they induce the release a sufficiently large amount of information about students at a single instant in time. Otherwise, the policy might not only be ineffective, but even exacerbate unraveling. Studying such policies is important in spite of the sizable literature focused with matching algorithms run by centralized clearinghouses, for two reasons. On the one hand, centralized clearinghouses are sometimes repudiated, as is the case in the market for law clerks. On the other hand, even markets with centralized clearinghouses sometimes unravel (Roth and Xing, 1994; Niederle et al., 2006).

To provide an empirical illustration of our model, we gather data from an internet discussion forum about the dates at which federal appellate judges hired law clerks for positions starting in the Fall of 2009.[3] Consistent with our model, hiring times are widely dispersed over a period of six months. To proxy for a judge's attractiveness as an employer, we use the number of students the judge has been able to place in Supreme Court clerkship positions between 1998 and 2008. As predicted, we find no relationship between this variable and the point in time at which the judge makes clerk hires.

A handful of previous papers model unraveling in two-sided matching markets. They are almost exclusively set in two periods. Specific to the market for law clerks, Haruvy et al. (2006) use laboratory experiments and computational methods to investigate proposed reforms in this market. The authors conclude that the feeling amongst students and judges that students cannot reject offers results in inefficiency in the market. Our model takes this restriction as a starting point. In contrast to both our and most other models of unraveling, Ostrovsky and Schwarz (2010) endogenize information revelation, and show that schools may prevent unraveling by disclosing just the right amount of information about students.

[3] Official data are not available. According to the Federal Law Clerk Hiring Plan which was in effect until 2013, judges were not supposed hire students before a set deadline. Hence, data about the points in time at which judges hire students would be a documentation of judge rulebreaking.

Other models each investigate a distinct aspect of unraveling. Halaburda (2010) shows that the likelihood of unraveling increases with respect to the similarity of participants' preferences. In Du and Livne (2013)'s model, approximately one quarter of participants have strict incentives to unravel as the market grows large, but this unraveling can be avoided by introducing transfer payments. Niederle et al. (2013) show how unraveling depends on imbalances in the demand and supply of applicants. Fainmesser (2013) studies a model in which information about market participants travels through a network of informal connections, and investigates the comparative statics of unraveling as a function of the structure of that network. Echenique and Pereyra (2016) precisely calculate the extent to which strategic unraveling pushes employers to hire early. Pan (2017) characterizes conditions on labor market characteristics and information structures that increase the chance of exploding offers.[4]

More generally, Doval (2017) and Kadam and Kotowski (2017) study stability notions in two-period models. Finally, a separate strand of the literature, starting with Li and Rosen (1998), studies unraveling in matching markets *with* transfers (Suen, 2000; Li and Suen, 2000, 2004; Du and Livne, 2013). In these models, early contracting may provide insurance to risk averse workers who face uncertainty about market conditions in the second period.

The remainder of this paper is organized as follows. Section 2 presents a brief overview of the market for law clerks as it relates to our model. Sections 3 sets up the model. Section 4 characterizes all equilibria of our model and analyzes comparative statics and welfare implications. Section 5 contains our empirical illustration of the hiring process for clerks to federal appellate judges. Section 6 studies policy proposals intended to alleviate unraveling. Section 7 formally compares our model to existing two-period models of unraveling. Finally, Section 8 concludes.[5] All mathematical proofs are in Appendix D.

2 The Market for Clerks to Federal Appellate Judges

Each year, students in law schools across the U.S. seek one of a small number of clerkship positions with a federal appellate judge. From the viewpoint of students, this is a highly elite

[4]Unlike these models, Damiano et al. (2005) set up a model in which unraveling is not associated with informational costs.

[5]Additionally, Appendix A presents an extension of the model in which students are allowed to reject offers, and an extension that shows the robustness regarding an alternative informational assumption. Appendix B is a numerical simulation of our model, and Appendix C describes our empirical data.

market. Federal appellate clerkships are the most prestigious of judicial clerkships available to new J.D. graduates, and are the toughest to obtain.[6] Nonetheless, some clerkships are more desirable than others. Judges differ in the extent to which they are involved in training clerks, so the choice of clerkship can have a significant impact on a clerk's further career development (Kozinski, 1991). Indeed, some judges have a reputation as *feeders* of clerks to the Supreme Court, and are regarded as particularly desirable employers.[7] Our empirical illustration in Section 5 relies on this fact to show the relationship between a judge's desirability as an employer and his or her clerk hiring time.

Also from the viewpoint of judges, finding good law clerks is vital. Having a talented staff of clerks is a major component of a their productivity, potentially making "the difference between a bad year and a wonderful one" (Kozinski, 1991). Judges thus compete fiercely for the best and brightest students in a cohort, often by asking students to sign a contract before they may consider offers from other judges.

As a result, the market for law clerks has suffered from unraveling for multiple decades. Both students and judges frequently complain about its chaotic and unpredictable nature. Repeated attempts to reform the market have invariably failed. Most recently, the Federal Law Clerk Hiring plan (or "Hiring Plan"), instituted in 2005 and officially abandoned in November 2013, was an attempt to coordinate the hiring process to begin no earlier than an agreed upon date (Online System for Clerkship Application and Review, 2013b).[8]

Our model matches three characteristics of this particular market, the first two of which are also present in many other entry-level labor markets. First, salaries are fixed and not negotiable. In the case of law clerks, this is stipulated by the Judicial Salary Plan.[9]

[6] Clerkship positions at federal appellate courts represent a small fraction of all judicial clerkship positions; about 10% of federal judges are at federal appellate courts (Administrative Office of the United States Courts, 2013).

[7] Many law school graduates view a clerkship at the Supreme Court as the ultimate achievement (Morse, 2014). These positions are more limited and more highly prized than clerkships with appellate judges. A lawyer who has committed to an appellate clerkship can command a signing bonus from law firms of up to $60,000 (Lat, 2011). Students transitioning from a role as Supreme Court clerk to an associate at a major law firm have been known to receive signing bonuses of up to $300,000 (Lat, 2013b).

[8] Adherence to the plan deteriorated as prestigious schools such as Harvard felt that such adherence placed students at a disadvantage and became lax in its enforcement (Lat, 2013a; Kramer, 2013).

[9] While fixed salaries imply that competitive pressure must be released through a channel other than transferrable utility, it is not true that allowing for flexible salaries would alleviate unraveling. There are many markets with flexible salaries that nonetheless unravel, such as the market for entry-level employees with J.D. degrees at elite law firms (Ginsburg and Wolf, 2003). For each recent cohort, there was a single salary ($160k in 2011) offered by the vast majority of elite law firms (Collins, 2012), rendering that market into one of *de facto* fixed salaries.

Second, information about students arrives gradually. Time to graduation in a J.D. program is three years and includes a large number of exams and evaluations, which progressively offer information about students. Judge Kozinski (1991) articulates the fundamental tradeoff: "Time ... can buy more information, reducing uncertainty. But the cost of delay may be the loss of opportunity, as some candidates receive offers from judges who believe they can make reasoned judgments on the basis of less information".[10]

Third, students feel pressured to accept any offer they might receive (Haruvy et al., 2006; Avery et al., 2001). Law career counselors reason that "one can never tell how far the wrath of a judge scorned will extend and exactly what the consequences will be, for you or for others" (Strauss, 2015). Indeed, the very act of turning down an offer may jeopardize chances for an offer from another judge. Strauss (2015) exhorts students to avoid declining any offer "merely as a matter of strategy with the expectation of receiving an offer from another judge".[11] Systematic survey evidence confirms the prevalence of this norm. In the year 2000, the most recent year for which such data are available, 73% of the students accepted the first offer they received (Avery et al., 2001).[12]

Reluctance to reject offers is arguably an idiosyncratic feature of the market for law clerks. Nonetheless, one can view it as an approximation to markets with a segment of *elite employers* if students are much more concerned about matching to one of the elite employers at all than to which of them they match to. In this case, they will be reluctant to refuse offers by elite employers, simply for strategic reasons.

3 Model Setup

In this section we introduce our formal model. A discussion of our assumptions follows.

Environment. We study a two-sided, one-to-one matching market without transfers. There are $J \geq 2$ employers and I students. The number of students exceeds the number of employers, $I > J$. Each employer can hire exactly one student. We think of $T > 0$ as the start date of employment. Each student i is identified with a skill level $q_i(t)$, given

[10] Kozinski (1991) lists specific pieces of information that, if available, are relevant to a judge's hiring decision.

[11] Additional evidence can be found in the recommendations of Yale Law School's Career Development Office, which highlight the extreme asymmetry of the judge-student relationship. Applicants are urged *not* to ask questions regarding the judge's policies on hours of work, vacation, or post-clerkship jobs lest they diminish their chances of receiving an offer (Yale Law School Career Development Office, 2008).

[12] There is evidence that students can be more strategic in their decisions which judges to apply to, and in accepting or avoiding phone calls (Avery et al., 2001). Our model will abstract from such behavior.

by an independent standard Brownian motion, with $q_i(0) = 0$ evolving over time interval $[0, T]$. Hence, predicting what a student's skill level will be at time T becomes easier as time progresses.

Preferences. Maximizing production is the employers' sole concern. Employer j's production is given by $v_j = \alpha_j q_i(T)$, with $\alpha_j > 0$ for all j. That is, employer j's production is proportional to the matched student's skill level at time T, $q_i(T)$. Hence, any employer who hires a student before the employment start date T needs to predict the expected skill level the student will attain by that point in time. The vector of parameters $(\alpha_j)_j$ allows us to capture complementarity or substitutability between students and employers. If α_j is larger for more desirable employers, student skill and employer desirability are complements. They are substitutes if α_j decreases with employer desirability. Employers not matched with a student receive payoff z, which we assume to be less than the ex-ante expected payoff that can be obtained from hiring the J-best student at time T. This assumption ensures that all employers will hire a student.

We assume that students have common preferences over employers. Specifically, each student prefers employer 1 to employer 2 to employer 3, and so on, and we say that employer 1 is *more desirable* than employer 2 who is in turn more desirable than employer 3, and so forth. While we frequently use language such as "least desirable employer" throughout this paper, we emphasize that this is a relative term in the context of a market consisting entirely of employers whose positions are so attractive that each student prefers them to outside options.

Actions and matching. At each time $t \in [0, T]$ employers must decide whether to remain in the market or exit the market by hiring a student. All matchings are binding for both sides. Students choose between employers if they receive multiple job offers at the same time. Otherwise, they must accept any job offer they receive.[13] The assumption of choosing between simultaneous offers is not vacuous since equilibrium incentives lead employers to act at the same points in time.[14]

Formally, at any moment in time $t \in [0, T]$ an employer can either stay in the market by choosing action s ("stay") or exit by choosing one of the actions in $E = \{e_1, \ldots, e_I\}$ ("exit" by employing a student). If, at time t, an employer is the only one to play some action $e \in E$, this employer is automatically assigned the student whose skill level is, at

[13] This assumption is arguably factually accurate in the market for law clerks; see Section 2.

[14] We study the effects of relaxing the no-rejections assumption in Appendix A.1.

time t, highest out of the students who have yet to sign an employment contract. If more than one employer attempts to exit at the same time t, students are matched to employers assortatively. An employer's choice of e_k indicates willingness to accept any student ranked k or higher amongst those who are still available, and prefers to stay in the game otherwise. Overall, this matching procedure embodies the assumption that a student may reject an offer by an employer if the student simultaneously receives an offer by a more desirable employer.[15]

Information. Employers' decisions are informed by the properties of the stochastic process according to which students' skill levels evolve. These are common knowledge, as are all players' past actions. Employers, however, do not observe the current realization of the student skill level $q_i(t)$ at any $t \leq T$. Employers can nonetheless infer the expected skill level obtained when they exit the game at any time t, conditional on the points in time $t_1 \leq \ldots \leq t_k$ at which previous employers have exited. This assumption renders the model a game of complete and perfect information.[16]

We use the expression $x_{k+1}(t|t_1, \ldots, t_k)$ to denote the skill level an employer expects from hiring the best available student at time t given the previous hire times of other employers $t_1, \ldots, t_k$. Observe that this notation allows us to capture the expected payoffs accruing to different employers even if they exit the market at the same point in time. For instance, if no employer has exited so far, and both employers 1 and employer 2 exit at t, then the more desirable employer 1 expects student skill level $x_1(t)$ and the less desirable employer 2 expects student skill level $x_2(t|t)$.

Histories, strategies, and equilibrium. A *history*, H_t, of the game up to time t describes what action has been taken by which player at what time up to, but excluding, time t. The history of the game is common knowledge. We let $\mathcal{H}$ denote the set of histories and we use the symbol $\emptyset$ to denote the empty history.

Intuitively, we think of a *strategy* as a plan that specifies the probability with which an employer takes a given action at any given point in time, depending on the actions previously taken by other employers. Formally, a pure strategy of an employer j is a function $\sigma_j : \mathcal{H} \to E \cup \{s\}$ that maps a history H_t into an action taken at time t. We write $\sigma_j(H_t) = a$ to refer to a strategy in which employer j plays action a at time t with probability 1. We

[15] Technically, this matching is derived from a serial dictatorship mechanism in which priority order is determined by the participating employers' desirability.

[16] We study the effects of relaxing this Assumption in appendix A.2.

require that any interval of time during which an employer plays some action $e \in E$ is closed on the left. This condition, which we call *right-continuity*, ensures that in any strategy profile there is a well-defined first point in time at which an employer exits the game. We emphasize that we *do* allow for mixed strategies. Since all equilibrium outcomes of interest can be obtained through pure strategy profiles, however, we defer the formalization of mixed strategies to Appendix D.

We derive our main result using subgame perfect Nash equilibrium as a solution concept, and also consider the limit of subgame perfect ϵ-equilibria for ϵ approaching 0.[17]

Discussion of assumptions

Interpretation of $q_i(t)$. Employers in our model are concerned with predicting the skill level that their prospective student hire will have on the start date of employment, T. Hence, our results are unchanged if we alter the process by which students' skill levels evolve in fully predictable ways, and in the same fashion for each student. For instance, our model can accommodate periods in time during which information arrives more quickly than usual (for instance the months during which many schools begin publicizing second year grades), as long as the release of a large amount of information is not concentrated on a single point in time.

Whether $q_i(t)$ represents information about skill that is learned in the moment, or whether it reflects gradually uncovered information about pre-existing differences in ability, is immaterial to our analysis. Similarly, $q_i(T)$ can be interpreted as a prediction of the student's expected productivity over the course of the match or, alternatively, as some objective measure of ability at that point in time, without any implications for the remaining elements of our analysis.

Imperfect information about $q_i(t)$. We assume that employers do not observe time t student skill levels $q_i(t)$ for reasons of tractability. Two further justifications are: First, our informational assumptions are a generalization of those in Halaburda (2010).[18] Second, we regard this assumption as plausible. If employers had perfect information, they would base their decisions not on the absolute skill levels, but on the difference between the skill levels of students of different ranks. It is not obvious that having perfect information about the

[17]For any $\epsilon > 0$, a strategy profile σ is an ϵ-equilibrium if no player can gain more than ϵ by unilaterally deviating from his strategy.

[18]In that model, I students who are undifferentiated in the first period are randomly assigned one of I fixed values in the second period. In our setup, students are undifferentiated at $t = 0$, expected skills conditional on rank j at time T are given by $x_j(T|T, \ldots, T)$, and ranks are assigned randomly.

difference in skill levels between two students would be a more plausible assumption than having no information other than the relative ranking of the students, especially if these two students are, for instance, the best student at Yale and the best student at Stanford.

Continuity of $q_i(t)$. In our model, information arrives continuously. Appendix A.3 shows that our substantive results are unchanged if information were to arrive in small discrete amounts. Analyzing the model as a sequence of discrete time games, one obtains, in the limit, the same set of equilibria as by directly using the continuous time formulation. Relatedly, in Sections 6.2 and 7 we study how the model changes when a *large* amount of information is released at a predictable point in time.

Student welfare. Because students in our model are ex-ante identical, they all face equal chances of matching to any given employer. Hence, their ex-ante expected welfare is independent of the way in which the market allocates students to employers.[19]

Student investment. Our model abstracts from human capital investment by students, which, in principle, might depend on the efficiency of the match and on the timing of the market. In the only related empirical study, using the market for college graduates in Japan, Okudaira (2017) finds that a guideline revision that successfully delayed the timing of job searches had no positive effect on students' human capital investment.

Employer desirability. All of our results continue to hold if higher ranking employers merely face a higher probability of being assigned better students when placed in direct competition with other employers.

4 Analysis

Our model is intuitive, and can be illustrated graphically, when there are only two employers. We discuss this case before we analyze the general model.

4.1 Intuitive Solution with Two Employers

Consider a market with two employers (let us call them employer 1 and employer 2), and at least four students.[20] Equilibria in our model are characterized by the points in time at which employers exit the market. Employers have no reason to exit the market early if there

[19]Note, however, from the time that students are differentiated, the rules of the market will affect the expected welfare of the students.

[20]This latter assumption ensures that t_1, as defined below, is strictly positive.

is no competition, but they do benefit from information that arises over time. Hence, once one of the two employers has exited the market, the remaining employer will remain in the market until time T. This employer can then hire the best of the remaining students with all pertinent information available. Thus, in any equilibrium, the second hire will be made at T. Therefore, we only need to characterize the point in time at which the first student is hired.

Figure 1 plots the time at which the first hire is made on the horizontal axis, and the student skill level an employer can expect to obtain from a particular strategy on the vertical axis. The employer who is first to hire a student and does so at t can expect the student to have skill level $x_1(t)$ at time T. This expectation is increasing in t: The closer t is to T, the higher is the likelihood that the student who was the most skilled at t will still be the most skilled at T.

The point in time t at which the first hire is made also affects the skill level $x_2(T|t)$ that the employer 2, who is second to hire, can expect at time T: The longer employer 1 stays before exiting, the greater is the chance that the student who is most highly skilled at T will be unavailable for employer 2. In this case, employer will have to settle for the student who is second-best at T. Hence, $x_2(T|t)$ is decreasing in t.

The two curves $x_1(t)$ and $x_2(T|t)$ cross. This is because at a very early points in time, it is difficult to predict which student will eventually be the most skilled. By hiring the student who is ranked first at such an early point in time, an employer cannot expect a payoff that is much higher than if he were to hire a randomly chosen student. If employer 1 were to do this, the probability is high that the student who eventually attains the highest skill level at time T will still available for employer 2. Hence, the strategy of waiting until T to hire the best remaining student yields a higher payoff than the strategy of making the first hire at very early points in time. Thus, for low t, we have $x_1(t) < x_2(T|t)$. On the other hand, if employer 1 waits until some late point in time to hire the first student (and employer 2 chooses not to preempt him by hiring first instead), employer 2 will likely have to settle for the student who is ranked second at T. Thus, for high values of t we have $x_1(t) > x_2(T|t)$.

Let t_1^* mark the point in time at which the two curves cross. Clearly, the first hire will not be made before t_1^*. At such early points in time, waiting beats exiting. On the other hand, there cannot be an equilibrium in which both employers postpone hiring until strictly after t_1^*. At such late points in time, exiting beats waiting. Hence, if an employer were

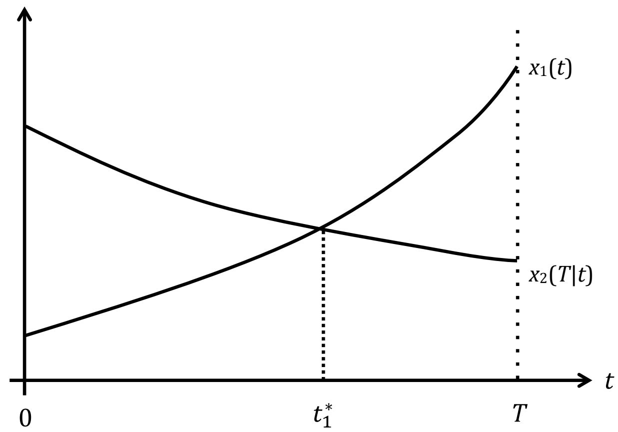

Figure 1: The model with two employers. The horizontal axis plots time t. $x_1(t)$ is the student skill level an employer expects to obtain if he hires a student at time t, given that no other student has been hired before. $x_2(T|t)$ is the student skill level an employer expects to obtain if he hires a student at T, given that one student has previously been hired at time t. In every equilibrium of the model, the first hire is made at t_1^*, and the second hire is made at T.

to employ a strategy of waiting until some $t > t_1^*$, the opponent could profitably preempt him. This leaves t_1^* as the only point in time at which the first student may be hired in equilibrium.

There *is* an equilibrium in which the first employer exits at t_1^* and the second employer waits until time T to exit. In this equilibrium, *both* employers attempt to exit by playing e_1 at t_1^*, and both employers receive the same expected payoff. No employer has an incentive to deviate: exiting before t_1^* is not profitable because at these early points in time, waiting beats exiting. Moreover, failure to play e_1 at t_1^* does not change either employer's expected payoff. Doing so will only cause one's opponent to exit at t_1^*, thus leaving the employer with payoff $x_2(T|t_1) = x_1(t_1)$. Finally, no employer has an incentive to deviate from remaining in the market until time T if the opponent was the first to hire at t_1^*, because the loss of information associated with exiting earlier will decrease the expected payoff.

Thus, unraveling proceeds to a point in time that is just early enough for both employers to expect to be matched to a student of the same skill level, regardless of whether they are first or second to exit the market. This is true in spite of the fact that both employers make an offer to the same student at the same point in time, and despite the fact that employer

1 is more desirable than employer 2. Our setup allows employers to avoid competition to such an extent that any advantage the more desirable employer might have when placed in direct competition with the less desirable employer is completely eroded.

In the equilibrium just discussed, the *more* desirable employer hires first. Yet, t_1^* is defined such that each employer gets the same expected payoff, regardless of whether the employer exits at t_1^* or at T. This raises the question of whether there is an equilibrium in which the *less* desirable employer exits first. In addition to presenting the full version of the model, the next section shows that under the appropriate equilibrium concept, our model indeed leaves the order of hiring undetermined.

4.2 Equilibrium Behavior

We now show that the intuition developed above extends to the case of arbitrarily many employers. We then consider the effects of unraveling on total production, and the comparative statics of unraveling with respect to the number of employers.

We begin the analysis by characterizing the employers' payoff functions, which are proportional to $x_{k+1}(t|t_1, \ldots, t_k)$. Recall that this expression denotes the expectation of the skill level that student i will have at the start date of employment T if i is the most highly skilled student at t amongst all students still available at t, given that at each previous time $t_1, \ldots, t_k$ an employer exited the market by hiring the respective best available student.

Proposition 1. *For all $0 \leq k \leq J-1$ and all $t \geq t_k \geq \ldots \geq t_1$, the conditional expectation $x_{k+1}(t|t_1, \ldots, t_k)$ is*

(i) jointly continuous in all arguments,

(ii) strictly increasing in t,

(iii) weakly decreasing in $t_1, \ldots, t_k$.

Moreover,

(iv) for all $t_1 \leq \ldots \leq t_J$ with $t_k < T$ for some k, we have $\sum_k x_k(t_k|t_1, \ldots, t_{k-1}) < \sum_k x_k(T|T, \ldots, T)$

(v) for all $0 \leq k \leq J-1$, $\tau_1 \leq \ldots \leq \tau_k$ with $t_1 > 0$, and $t \geq t_k$, we have $x_{k+1}(t|t_1, \ldots, t_k) < x_k(t|t_1, \ldots, t_{k-1})$

Property (i) captures the intuition that information arrives gradually. It does not hold for a process in which students' skill levels jump with positive probability at some known or predetermined time. Property (ii) arises because predicting farther into the future is more difficult. The closer t is to the start date of employment, T, the higher the probability that the student who is deemed best at time t will still be ranked highly at T. As a natural analogue, we obtain property (iii). The earlier an employer's rivals hire, the more likely the hired students have dropped in rank by time t, and hence the better the pool of remaining students from which the employer can choose. Property (iv) states that unraveling is inefficient from the employers' point of view. If hirings are made early, then there is a positive probability that some employer will be matched to a student whose rank at T is lower than J. Such a student would not be hired at all if all employers postponed hiring until T. Finally, the expected payoff of an employer who exits at time t decreases with the number of students that have already matched to other employers. This intuition is formalized in property (v). It is only through these properties that all of our subsequent results depend on the process of student skill evolution. Hence they hold for any stochastic process for which Proposition 1 can be derived.

Within each subgame, equilibrium strategies are characterized by the points in time at which employers will rationally exit the market. For each history H_t in which k employers have exited, the *subgame hiring times* $t_{k+1}(H_t), \ldots, t_J(H_t)$ are a set of points in time at which employers will weakly prefer exiting to staying in the game (given that other employers plan to do the same). They are defined by four properties.

First, the sole reason to hire before T is competition with other employers. Hence, if an employer finds himself to be the only employer who has not yet hired a student, the employer will wait to do so until time T when all pertinent information is available. Second, exiting at a later subgame hiring time must not be more profitable than exiting at an earlier subgame hiring time, as employers would then have an incentive to postpone hiring. Third, if all future exits from the market occur at subgame hiring times, each employer is indifferent between exiting at any of the (strictly) future subgame hiring times. Fourth, whenever immediately hiring any of the m best available students is more profitable than remaining in the game, many employers will attempt to exit the market immediately, resulting in a *hiring frenzy*. Hiring frenzies may never be known to occur in the (strict) future: if they were, less desirable employers would have an incentive to hire the best available student just before the hiring frenzy is known to take place.

The following definition states these properties formally.

Definition 1. *(Subgame hiring times) Consider a history H_t with past hiring times $h = (t_1, \ldots, t_k)$ that satisfies $t_1 \leq \ldots \leq t_k \leq t$. Then $t_{k+1} \leq \ldots \leq t_J$ are* subgame hiring times *for H_t if they satisfy $t_{k+1} \geq t$ and*

(i) $t_J = T$.

(ii) For all $j \in \{k+1, \ldots, J-1\}$:
$x_j(t_j|h, t_{k+1}, \ldots, t_{j-1}) \geq x_{j+1}(t_{j+1}|h, t_{k+1}, \ldots, t_j)$.

(iii) If $t < t_j < t_{j+1}$ for some $j \in \{k+1, \ldots, J-1\}$, then $x_j(t_j|h, t_{k+1}, \ldots, t_{j-1}) = x_{j+1}(t_{j+1}|h, t_{k+1}, \ldots, t_j)$.

(iv) If $t_j = t_{j+1}$ for some $j \in \{k+1, \ldots, J-1\}$, then $t_{j+1} = t$.

The set of equilibria in our game is small because the monotonicity of the conditional expectation functions guarantees unique subgame hiring times for each history.

Lemma 1. *For all $k \in \{0, \ldots, J-1\}$ and for each history H_t in which k employers have exited, there exists a unique vector of future subgame hiring times $t_{k+1}(H_t) \leq \ldots \leq t_J(H_t)$.*

We can now explicitly construct a pure strategy equilibrium profile σ. Employer j's strategy in this profile, σ_j, takes the following form.

Definition 2. *(Equilibrium Strategies) A strategy σ_j for player j is an* equilibrium strategy *if, for each t and for each history H_t with previous hiring times $h = (t_1, \ldots, t_k)$ that satisfy $t_1 \leq \ldots \leq t_k \leq t$, employer j's strategy satisfies*

(i) $\sigma_j(H_t) = s$ *if* $t < t_{k+1}(H_t)$.

(ii) $\sigma_j(H_T) = e_J$.

(iii) $\sigma_j(H_t) = e_m$ *if* $t_{k+1}(H_t) = t_{k+m}(H_t) = t$ *and if*
$x(t_{k+m}(H_t)|h, t_{k+1}(H_t), \ldots, t_{k+m-1}(H_t)) > x(T|h, t_{k+1}(H_t), \ldots, t_{J-1}(H_t))$.

(iv) $\sigma_j(H_t) = s$ *or* $\sigma_j(H_t) = e_1$ *at any other subgame hiring time.*

Part (i) of this definition ensures that no employer exits at a time that is not a subgame hiring time. Part (ii) ensures that employer j agrees to hire any student who is at least J-best at T. Part (iii) requires that employer j attempts to exit the game whenever doing so is strictly more profitable than remaining in the market (and thus participates in every hiring frenzy in which that employer has a chance of employing a student). It also ensures that at least one employer exits at each subgame hiring time: any attempt to postpone hiring by $\epsilon > 0$ will be preempted by some other employer. Finally, by part (iv), employer j may or may not attempt to hire a student at any other subgame hiring time.

The strategy profile σ is an equilibrium profile if the ambiguity in part (iv) is resolved as follows: At each time t that coincides with $l \geq 1$ subgame hiring times, the $(l + 1)$ most desirable remaining employers play action e_l. The $(l + 1)$st employer ensures that none of the l employers who do exit have an incentive to postpone exiting by any (arbitrarily small) amount of time.

The next proposition is our main result. Part (i) shows that σ is an subgame perfect equilibrium (SPE) profile. Part (ii) shows that, in each SPE profile and each subgame, students are hired at subgame hiring times. Hence, any equilibrium outcome of the game has one employer exit at each subgame hiring time $t_1(\emptyset), \ldots, t_J(\emptyset)$. We call these points in time the *equilibrium hiring times*, and denote them by $t_1^*, \ldots, t_J^*$.

It bears emphasis that proposition 2 holds true even though we allow players to play mixed strategies.

Proposition 2.

(i) σ is an SPE profile.

(ii) In each SPE profile, students are hired at subgame hiring times in each subgame.

Testable implications. Proposition 2 implies that equilibrium hiring times will be spread out.[21] The extent of the dispersion in equilibrium hiring times is substantial: the first employer exits the market strictly before the point in time at which hiring the best

[21] Under the assumptions of Section 3 no hiring frenzy will occur in equilibrium. A hiring frenzy may occur, however, if a large amount of information becomes available at a predictable point in time. This is the case, for instance, if a policy alters the release of information about students (see Section 6.2), or if students are already sufficiently differentiated at $t = 0$.

student would provide him with an expected payoff as large as randomly matching to one of the J most highly skilled students at T.[22]

Additionally, all employers obtain a student with the same expected skill level, irrespective of their employer desirability rank (unless there is a hiring frenzy at $t = 0$). This drives our result regarding the order in which employers exit the market. Since each employer is indifferent between exiting at any equilibrium hiring time, any order of exit is consistent with a limit of subgame perfect pure strategy ϵ-equilibria, for ϵ approaching zero. (By contrast, if we reduce our model to two periods, and if the first period features little information about students, as is the case in Halaburda (2010), we find that less desirable employers are more likely than more desirable employers to hire in the first period. See Section 7.)

An important caveat to the above result is that the only order of exit consistent with exact equilibrium has the most desirable employer exit first, followed by the second-most desirable employer, and so on. This is an implication of the right-continuity of strategies. By part (iii) of definition 2, an employer must attempt to exit the game whenever doing so is strictly more profitable than the continuation strategy. By right-continuity, therefore, each employer must also exit when doing so is weakly more profitable. Thus, at each point in time, the most desirable remaining employer must exit at the next subgame hiring time.[23] Because this implication on the order of exit derives from a purely technical condition (right-continuity) rather than from a substantive assumption, we regard the anything-goes result implied by limits of ϵ-equilibria as the main prediction of our model regarding the order of exit. To demonstrate that our results are not an artifact of our particular continuous-time formulation, Appendix A.3 re-derives the results using discrete time approximations for the case of two employers.

Lastly, we find that increased competitive pressure exacerbates unraveling. If a larger number of employers compete for the same set of students, then the jth student is hired at an earlier point in time for all j.

Corollary 1. *If there is no hiring frenzy at $t = 0$, then the following hold:*

(i) Equilibrium hiring times are dispersed in time: $t_1^* < \ldots < t_J^*$.

[22]This follows from the fact that the first student is hired as soon as doing so yields the equilibrium expected student skill level, and because the equilibrium expected student skill level is lower than that from the random matching.

[23]ϵ-equilibrium allows us to circumvent this implication of right-continuity. A more desirable employer nonetheless has a competitive advantage over less desirable one. Whenever desired, a more desirable employer can exit at the same point in time as a less desirable employer, and thus preempt him.

(ii) In any equilibrium, all employers obtain a student of the same expected skill level.

(iii) Any order of employers exiting the market is consistent with a limit of subgame perfect pure strategy ϵ-equilibria, for ϵ approaching zero. The unique order of employers exiting the market consistent with exact equilibrium has the j-most desirable employer be the jth to exit.

(iv) Consider an increase in the number of employers to $\tilde{J} > J$. Let $\tilde{t}^*_1, \ldots, \tilde{t}^*_{\tilde{J}}$ denote the associated equilibrium hiring times. Then, $\tilde{t}^*_j < t^*_j$ for all $j = 1, \ldots, J$.

Distributional implications and total production In many markets, unraveling is regarded as an undesirable state of affairs, and is often compared to a prisoner's dilemma. That comparison, however, is misleading, because unraveling is not Pareto inferior to a positive assortative matching at T — the unique *ex post* pairwise stable match in our setting. When employers and students are matched assortatively, the least desirable employer matches to a student with expected skill level $x_J(T|T, \ldots, T)$. In the unraveling equilibrium, the least desirable employer exits last, and because $x_J(T|t_1, \ldots, t_{J-1})$ is decreasing over the previous hiring times $t_1, \ldots, t_{J-1}$, this least desirable employer obtains a strictly higher payoff in an unravelled market than under assortative matching.

It is true that unraveling results in lower expected student skill levels for employers who hire at T. Whether or not unraveling decreases the total output of the market, however, would also depend on whether employer desirability and student skill are complements or substitutes. While total production in the unraveling equilibrium does not depend on substitutability,[24] the assortative matching benchmark does. Positive assortative matching maximizes production in the complements case, and minimizes it in the substitutes case. With substitutes, any deviation from positive assortative matching increases production (keeping the set of matched students unchanged); if substitutability is sufficiently strong (that is, if α_j increases with j sufficiently quickly.) this effect outweighs the loss in expected student productivity caused by unraveling.[25]

Corollary 2.

[24]Expected total production in any unraveling equilibrium is given by $E\left[\sum_j v_j(q_{\mu(j)}(T))\right] = E\left[q_{\mu(1)}(T)\right]\sum_j a_j$, where $\mu(j)$ denotes the index of the student that employer j is matched to in a given realization of an equilibrium.

[25]This result crucially depends on the absence of transfers.

(i) Less desirable employers strictly prefer any unraveling equilibrium to a positive assortative matching at T.

(ii) Production in positive assortative matching at T is higher than in any unraveling equilibrium if student skills and employer desirability are complements, and lower if they are sufficiently strong substitutes.

Whether the substitutes or the complements case better describes the market for law clerks is an open empirical question.[26]

5 Data

We now present data from the 2008 market for law clerks. Consistent with our model, hirings are indeed spread over time, and we find no relationship between judges' desirability and the point in time at which they hire students. We do not, however, interpret these findings as a *test* of our model. The Hiring Plan was still in effect in 2008, and thus our data may not represent equilibrium behavior.

Data We source our data from the "Clerkship Notification Blog", a public internet forum coordinated by law students (McDaniel et al., 2013), the central objective of which is to divulge information about the timing of judicial clerkship hirings and interviews.[27] It largely consists of time-stamped entries such as "Ikuta has hired one HLS" or "Noonan is done".[28] While blog entries are anonymous, they are often seconded and mistakes based on hearsay are frequently corrected. Often, they are written by current law clerks divulging inside information.

Our data covers the hiring for positions starting in 2009, and contains information about the hiring activity of 197 circuit judges. We restrict our data to statements of a judge hiring a clerk or finishing the hiring process. (A judge may hire several clerks in a given year.) This leaves us with 168 judges, which covers just over 80% of the 239 judges currently active

[26]The case of substitutes, however, is not merely a theoretical curiosity. Davis (2017) examines the matching of teachers to schools with Teach For America. He examines the change in the matching procedure from a first-offer mechanism, which induced some unraveling, to a deferred-acceptance algorithm. He finds that while teacher retention increased (indicating that the new match outcomes better account for teacher preferences), it lowers performance in the teachers' first year.

[27]Official records on the points in time at which judges hire clerks do not exist, since judges who hire earlier than the Hiring Plan date are breaking the very rule this plan was meant to uphold.

[28]Meaning "Judge Ikuta has hired one Harvard Law Student", and "Judge Noonan has finished hiring for this season".

in the districts that we analyze. We merge the timing data with data about the number of clerks a judge has sent to Supreme Court clerkships between 1998 and 2008, our proxy for the desirability of judges. As in the case for timing data, there is no official data. Instead, we use the user-compiled dataset on Wikipedia (2013), which lists more than 1900 law clerks to the Supreme Court, and the clerkship held previously. Details about data collection and encoding are in Appendix C.

Analysis Figure 2 plots, for each judge, the first and last points in time an entry was posted about a judge hiring a clerk or completed the hiring process. The dashed vertical line indicates the hiring plan date. We observe significant unraveling. Hires are spread over a time period of about half a year, and occur up to five months before the Hiring Plan date. Strikingly, at almost no point in time between April and December of 2008 was there a lull in hiring, and the only point in time at which hirings bunch is the Hiring Plan date (as can be seen by the much steeper increase in the slope in Figure 2 shortly after that date).[29]

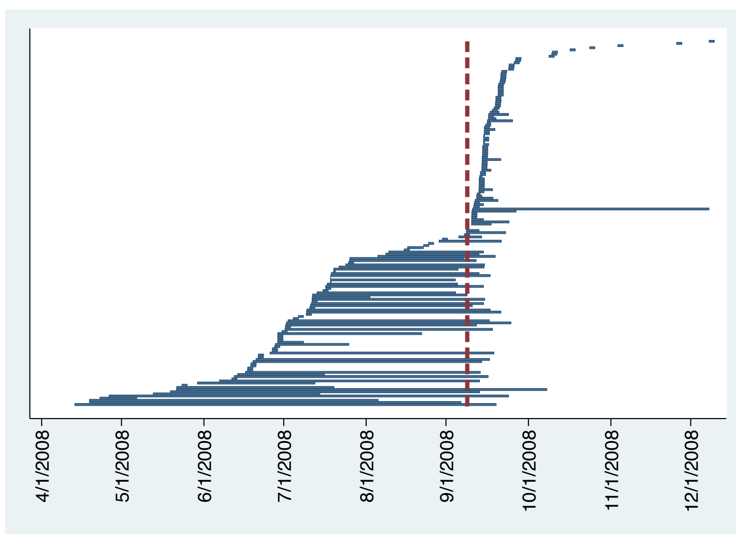

Figure 2: First and last hiring or finish data points by judge.

Our model links hiring times to the extent of information that is available about students. The majority of law schools finalized second year grades by the end of June.[30] Hence, the slight increase in activity we observe around June is consistent with our model. Notably,

[29]Note that the left ends of the bars in figure 2 define an empirical cumulative distribution function.
[30]Online System for Clerkship Application and Review (2013a).

judges hiring significantly before this date do so at a loss of information. Information about students likely continues to arrive even after the release of grades, since many students take their first position as an intern at a law firm in the summer between second and third year. Thus, our model accounts for the continued hiring activity in July and August. Overall, the dispersion of hiring times in this market is consistent with the predictions of our model. It is reminiscent of the substantial spreading out of offer times that Niederle et al. (2006) document in the market for gastroenterology fellowships.[31]

While our data cover one single year, dispersion in hiring times appears to persist past 2008. Tobias (2013) details the various states of hiring that different courts are in for October 2013, which fall into six categories ranging from "finished hiring" to "have recently started or not yet begun" and "remain unclear". A small number of blog postings from the hiring seasons in 2013 through 2015 are available on a different forum[32] and suggest that hirings are spread over similarly long intervals of time.[33]

We now turn to the relationship between the time at which judges hire and the desirability of the positions they offer. Figure 3 plots, for each judge, the number of clerks sent to the Supreme Court between 1998 and 2008 against the point in time at which we first observe the judge in our data. While the more desirable judges avoid hiring substantially after the Hiring Plan date, there is no relationship between the number of clerks sent to the Supreme Court and the hiring time for those judges who hired before or around the Hiring Plan date.[34]

Formally, we regress the number of clerks sent to the Supreme Court on the earliest hiring time from the Clerkship Notification Blog data. Using the entire sample, we find that a judge who begins hiring 30 days later, ceteris paribus, has sent an average of 0.41 fewer students to the Supreme Court between 1998 and 2008 ($p = 0.075$). This number changes to 0.69 ($p = 0.218$) once we include only judges who hired at or before the Hiring Plan date, September 8, and to 0.33 ($p = 0.303$) once we include observations up to 5 days after the Hiring Plan date (allowing for delays in reporting). We conclude that there is no clear

[31] That data was gathered in the aftermath of a disruption of that market, and thus may not reflect equilibrium behavior.

[32] www.top-law-schools.com (2015)

[33] For the 2014 season, www.top-law-schools.com contains reports from January 2013 that judge Kozinski has finished hiring, and in August 2013 that judge Hamilton has finished hiring. For the 2015 season, two judges are reported to have finished hiring in August 2013, one judge is reported to be hiring in April 2014, and others have unclear status in October 2014.

[34] Our model does not incorporate any effects of 'soft' interventions such as the Hiring Plan, and thus makes no predictions about behavior close to that date.

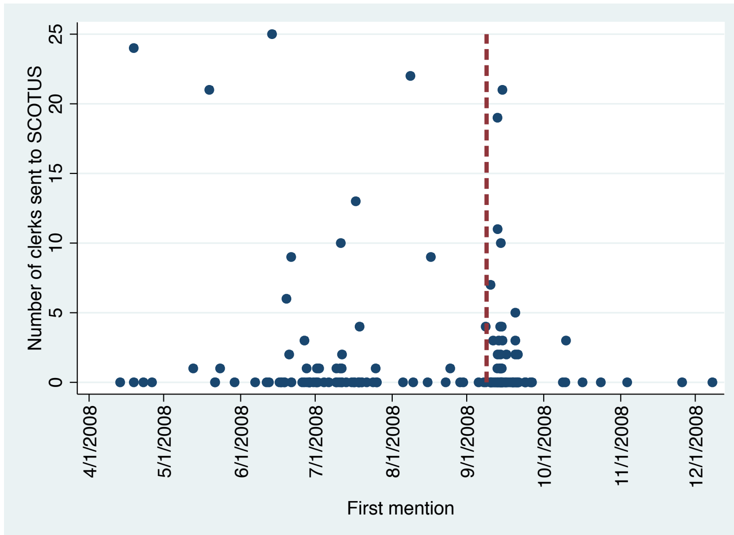

Figure 3: Number of clerks sent to the Supreme Court by a judge between 1998 and 2008 versus earliest hire.

relationship between the first time a judge hires a student and the judge's desirability, as predicted by our model.[35]

6 Policy Implications for the Law Clerks Market

Our model suggests that there is no reason to hope that 'soft' interventions such as appeals to wait for a longer time before hiring students may lead to coordination on some 'better' equilibrium in the market for law clerks. All equilibria in our model give rise to the same set of hiring times and the same expected payoffs for employers. Hence, a policy can have lasting effects only if it changes either payoffs associated with different strategies, or the information that is available to judges. This may explain why all past attempts at coordinating hiring to a specified point in time have failed.

6.1 Clearinghouses and Matching Algorithms

Instituting a centralized clearinghouse run at time T is an often discussed proposal that alters payoffs associated with different strategies. In our model, the unique pairwise stable match

[35]There are two reasons for a judge's superior ability to send clerks to the supreme court. First, the judge may be able to provide more value added to her current clerks than other judges, or have better connections to the Supreme Court. Second, a judge may have a superior ability to select good students. Judge Wald (1990) appreciates both channels, Judge Kozinski (1991) emphasizes the former. The data confirm the predictions of our model to the extent that differences in judges are due to the former channel.

would result if all employers postponed hiring until the arrival of all pertinent information. Thus, it is perhaps unsurprising that the equilibrium outcome in our model is entirely unaffected by such a change, regardless of the matching mechanism used.[36] The reason is intuitive. If two or more employers participate in a clearinghouse run at time T, it is necessarily the case that one of these employers faces a positive probability of not matching with the best available student. This employer can jump the gun and hire the best available student slightly before the start of the clearinghouse, reaping a discrete gain in expected student skill level. Hence, at most one employer would ever be willing to participate in such a clearinghouse.[37]

Corollary 3. *A centralized clearinghouse run at T does not alter the set of equilibrium outcomes regardless of the matching mechanism employed.*

This result is due to the gradual arrival of information rather than due to our assumption that students are unable to reject offers. To see this, consider our model with the sole change that students *may* reject offers, and suppose that the market unravels. In our setting, assortative matching is the *unique* stable match. Hence, for any other mechanism, there will be a blocking pair, that is, a student and an employer who would prefer matching to each other over their respective mechanism-assigned match. Because information arrives continuously, this pair would be able to find each other with very high probability and sign a contract at some time slightly before T, thus blocking the match.[38] Hence, alternative mechanisms cannot prevent unraveling. This contrasts with the two-period model in Halaburda (2010) in which there always exists a (pairwise *un*stable) mechanism that prevents unraveling.

[36] Roth and Xing (1994) and Kagel and Roth (2000) argue that whether unraveling occurs in a two-sided matching market with a centralized clearinghouse depends on whether the mechanism it employs is pairwise stable. While a stable matching procedure appears necessary to prevent unraveling, it is not sufficient, as the current paper shows theoretically. Empirically, the Canadian market for lawyers exhibits severe unraveling even though a centralized clearinghouse is in place (Roth and Xing, 1994). Additionally, Wetz et al. (2010) study the market for hospital interns in the year 2007, which was cleared using a pairwise stable matching algorithm, and find that 15.7% of the postgraduate positions were offered outside the match.

[37] Related to this result, Haruvy et al. (2006) conclude that unraveling in the market for law clerks is not due to lack of pairwise stability, but due to the ease with which binding contracts are forged.

[38] Formally, if employer j and the student ranked i at time T are a blocking pair, then the probability that the student who is ranked i at $T-\eta$ will still be ranked i at T is arbitrarily close to 1 for sufficiently small η. Consequently, if employer j and the student who is ranked i at $T-\eta$ match at $T-\eta$, the probability that one of these agents would prefer to have participated in the matching mechanism at T is arbitrarily small.

6.2 Changes in the Arrival of Information

In contrast to a centralized clearinghouse, a policy intervention that alters the process by which information about students is divulged may affect hiring behavior in our model. The effectiveness of such a policy crucially depends on the details of the implementation.

Continuous Changes in the Arrival of Information. A continuous change in the arrival of information will only affect the points in time at which hires are made, but not the equilibrium expected payoff of any market participant.

Formally, we define a change in the arrival of information as follows. Let $\gamma : [0, T] \to [0, T]$ be a strictly increasing function with $\gamma(0) = 0$ and $\gamma(T) = T$. Information that arises at t is made available to employers at $\gamma(t)$. Hence, if an employer plays e_1 at time $\gamma(t)$, this employer is assigned the highest ranked available student at t.

Corollary 4. *If $\gamma : [0, T] \to [0, T]$ is continuous, then in any subgame perfect equilibrium of our model,*

(i) students are hired at times $\gamma(t_1^*), \ldots, \gamma(t_J^*)$

(ii) the equilibrium expected utility of each market participant is the same as without the policy.

This corollary is immediate. Continuity of γ ensures that the conditional expectation functions are continuous. Therefore we can construct the set of equilibria just as before, by finding points in time where the expected payoff from exiting the market equals the expected payoff from staying in the market. Importantly, equilibrium hiring times are defined not by how much time is left until employment starts, but by the amount of information present at a given point in time. Since it is the latter that determines employers' utility in the game, the policy has no effect on their production. Moreover, both with and without the policy, all students are identical at time $t = 0$, and thus face the same distribution over eventual outcomes. Therefore, students' *ex-ante* expected utilities are unaffected by the policy change.

Information Discontinuities. Employers' equilibrium payoffs *can* be affected if the information revelation policy introduces a discontinuity in the conditional expectation functions. It is crucial that the discontinuity is sufficiently large, as the policy may otherwise exacerbate unraveling.

We study the case in which information that arrives in some interval of time $[\underline{t}, \bar{t})$, with $\underline{t} < \bar{t}$, is withheld, and discontinuously released at $\bar{t}$. Formally, $\gamma(t) = t$ if $t \in [\underline{t}, \bar{t})$, and $\gamma(t) = \underline{t}$ if $t \notin [\underline{t}, \bar{t})$. Hence, if some employers play actions $e \in E$ during the interval $[\underline{t}, \bar{t})$, they are assigned the students they would have been assigned if they all had played the these actions at time $\underline{t}$—except that it is the earlier judges rather than the more desirable judges who are then allocated the better students. Therefore, any employer who intends to exit at some time $t \in (\underline{t}, \bar{t})$ can do at least as well by exiting at $\underline{t}$. Thus, this policy is equivalent to preventing exit in the time interval $(\underline{t}, \bar{t})$.

The effect of the policy is best explained in the case of two employers. Plainly, the policy does not eliminate any equilibria if $[\underline{t}, \bar{t})$ does not contain the equilibrium hiring time t_1^*. In all other cases, the location of $\underline{t}$ and $\bar{t}$ determines whether the policy alleviates or exacerbates unraveling.

Consider the graphs in Figure 4, which parallel those of Figure 1. In both panels, information is released at the same time $\bar{t}$. If neither employer hires a student before $\bar{t}$, then exiting at $\bar{t}$ is strictly better than remaining in the market. Thus, both employers will attempt to exit at this time. The more desirable employer 1 will successfully do so, obtaining a student of expected skill level $x_1(\bar{t})$, while the less desirable employer 2 will hire the best remaining student at T, expecting a student with skill level $x_2(T|\bar{t}) < x_1(\bar{t})$. Whether employer 2 can do better by exiting before $\bar{t}$ depends on the start time $\underline{t}$ of the withholding period.

In panel (A), the withholding of information starts so early that $x_1(\underline{t}) < x_2(T|\bar{t})$. Hence, employer 2 prefers hiring the best remaining student at T to exiting the market at any time at or before $\bar{t}$. In this case, the policy alleviates unraveling. In panel (B), the withholding of information starts late, so that $x_1(\underline{t}) > x_2(T|\bar{t})$. In this case, employer 2 prefers exiting at $\underline{t}$ to hiring the best of the remaining students at T. Since $\underline{t} < t_1$, the policy exacerbates unraveling.

Intuitively, the withholding policy has the potential to exacerbate unraveling because it places employers in direct competition with each other at time $\bar{t}$ when information is released. While more desirable employers benefit from the policy, less desirable employers lose. Therefore, the policy provides additional incentives for less desirable employers to hire early.

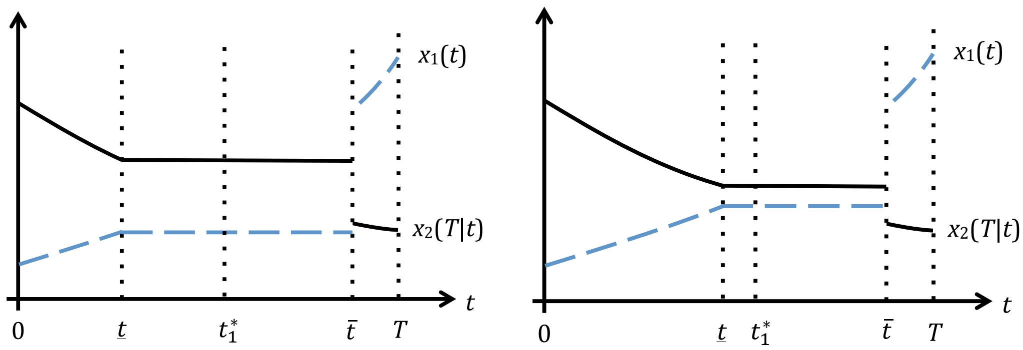

Figure 4: The effects of withholding information about students in the case of two employers. In panel (A), the withholding of information starts early so that the first student is hired at $\bar{t}$. In panel (B), the withholding of information starts late. The first student is hired at $\underline{t}$ and so the policy exacerbates unraveling.

The following proposition formalizes this discussion for the case of J employers. The condition $0 < x_J(T|T, \ldots, T)$ in part (i) is necessary, as otherwise the least desirable employer prefers hiring at 0 to hiring at T even if information is withheld in the interval $[0, T)$.[39]

Corollary 5. *Suppose that information about students is withheld on the interval $[\underline{t}, \bar{t})$ and discontinuously released at $\bar{t}$. Then the following holds:*

(i) If $\underline{t}$ is sufficiently low and $0 < x_J(T|T, \ldots, T)$, then no students are hired before $\bar{t}$. Students are hired at the subgame hiring times for history $H_{\bar{t}}$ in which no students are hired during the time interval $[0, \bar{t})$.

(ii) Withholding information can exacerbate unraveling: there are values $\underline{t}$ and $\bar{t}$ with $\underline{t} < \bar{t}$ such that the first student is hired strictly before t_1.

A policy of withholding information about students successfully curtailed unraveling in the market for medical residency internships in 1945 (Roth, 1984). The policy was implemented boldly. Medical colleges agreed to release neither academic transcripts nor letters of recommendation before an agreed-upon date.[40] In other markets, it may be

[39] This is the case, for instance, if the number of students is smaller than twice the number of judges, i.e. $I < 2J$. In this case, the expected skill level of the J-best student at T is negative, and the least desirable employer would be better off hiring a random student at $t = 0$.

[40] The matching that resulted after the release of information was, however, not pairwise stable, which led to a different set of issues that ultimately led to the adoption of what is known today as the National Resident Matching Program. See Roth (1984), and Roth and Peranson (1999).

practically difficult to withhold information for the required amount of time, and to prevent leakage of information. Indeed, such an attempt faced very limited success in the market for law clerks (Wald, 1990).

Withholding information about students is not the only means of generating an informational discontinuity. One alternative consists of limiting the production of information about students, for instance by switching from letter grades to a pass/fail system.[41] Another possibility is a nationwide test that is designed to provide a sudden surge in information about students' abilities. An informational discontinuity is introduced by releasing the test results at a pre-specified point in time.

Regardless of the specific policy used, a hiring frenzy should be expected when information is released. In order to prevent market congestion, it may be necessary to employ a centralized clearinghouse at that time.[42]

7 Relation to Models of Early Offers

Most existing work on unraveling in matching markets restricts market participants to transact at one of two points in time. Imposing such a restriction in our model substantially changes equilibrium outcomes, as it allows employers to avoid competition to a much more limited extent. Thus, unraveling only partially erodes the advantages of the more desirable employers, and hence leads to less redistribution in the restricted model than in the unrestricted one.

The specific properties of the restricted model depend on the amount of information that is available in the first period. If there is very little information, it resembles the model in Halaburda (2010). In this case, pure strategy equilibria exist, and only less desirable employers hire in the first period. This latter prediction contrasts with both the unrestricted model and the empirical evidence in Section 5. Intuitively it arises because with very little information in the first period, every employer obtains a student with expected skill level close to zero from hiring early. In the second period, more desirable employers manage to match to the better students. Given this, if there is some employer who finds it profitable

[41] Yale Law School has used pass/fail grading for first year students in those years in which unraveling was particularly extreme (Al Roth, personal communication). Ostrovsky and Schwarz (2010) show how limiting information about students in a particular fashion may prevent unraveling.

[42] The more positively assortative the matching algorithm used by the clearinghouse, the smaller the least desirable employer's expected payoff from participating in the mechanism, and hence, the larger the discontinuity needs to be to prevent this employer from hiring early.

to avoid competition with strong employers in the second period and instead obtain a student with expected skill level close to zero, then any less desirable employer will find this profitable, too. By the same argument, we see that expected equilibrium payoffs are increasing in employer desirability.

If the amount of information available in the first period is sizable, our restricted model resembles the model in Du and Livne (2013). The game then has a matching-pennies structure: Less desirable employers prefer to avoid competition with stronger employers, but stronger employers prefer to directly compete with less desirable ones.[43] Hence, any equilibrium will be in mixed strategies. If the extent of information available in the first period is sufficiently large, then the expected equilibrium payoffs are increasing in employer desirability in any mixed strategy equilibrium.

Proposition 3. *Let* $t_0 \in [0, T)$ *and suppose that employers can exit the market only at times* t_0 *and* T*. Then,*

(i) if $t_0 = 0$*, there exists a pure strategy equilibrium profile, and every such profile is in threshold strategies: there is an employer* $\hat{j} \in \{1, \ldots, J\}$ *such that employers* $1, \ldots, \hat{j}$ *exit at* T *and all other employers exit at* t_0*. In this case,* $\pi_1 > \ldots > \pi_{\hat{j}-1} \geq \pi_{\hat{j}} = \ldots = \pi_J$*. If, moreover,* $x_J(T|T, \ldots, T) > 0$*, then the unique equilibrium is for all employers to exit in the second period.*

(ii) If t_0 *is sufficiently close to* T*, then no pure strategy equilibrium exists, and* $\pi_1 > \ldots > \pi_J$ *in any mixed strategy equilibrium.*

Two further differences between the restricted and unrestricted versions of our model are noteworthy. First, if the restricted model features little information in the first period and if $x_J(T|T, \ldots, T) > 0$,[44] then there will be no unraveling. While less desirable employers would like to avoid competition with stronger employers, they could do so only at an unprofitably large informational cost. In contrast, in the unrestricted model, less desirable employers can always avoid competition by hiring just slightly before stronger employers, at an arbitrarily small informational cost. Hence, the unrestricted model always features unraveling, and the vast majority of positions are filled before all pertinent information about students is available.

[43]If weaker employers hire in the first period, stronger employers will also prefer hiring in the first period in order to avoid being preempted, whereas if weaker employers hire in the second period, stronger employers prefer hiring in the second period in order to reap the informational gains from postponing hiring.

[44]This is the case if $I > 2J$.

Second, while matching mechanisms are of limited usefulness in the unrestricted model (see Section 6.1), there *always* exists a (pairwise *un*stable) second-period matching mechanism that completely eliminates unraveling in the restricted model with $t_0 = 0$.[45] The existence of such a mechanism relies on the fact that employers in the restricted model have far fewer opportunities to avoid the clearinghouse, and that they can do so only at a large informational cost.

8 Conclusion

Unraveling is a multifaceted issue that plagues many entry-level labor markets. We complement the literature by providing a model that focuses on the strategic timing aspect of the phenomenon. Our model is inspired by the market for law clerks. It thus requires students to accept any offer they receive, unless they receive multiple simultaneous offers, in which case they may choose the most preferred one. Employers pick students from the market as soon as doing so yields the same expected payoff as the continuation strategy. This behavior thins out the market as it leads to widely dispersed hiring times, and it perfectly equalizes the student skill level that each employer expects to obtain in equilibrium. This is true even if all (but the last) hired students receive multiple simultaneous offers. The primitives of our model – employer desirability and student skill levels – do not tie down the relation between an employer's desirability and the point in time at which this employer hires a student. We expect that this relation is determined by factors orthogonal to our model.

We gathered data on the hiring times of federal appellate judges in 2008, and found them to be dispersed over a period of six months. Using the number of clerks a judge has been able to place at the Supreme Court as a proxy for law students' preferences over judges, we confirm that there is no statistically significant relationship between the point in time at which a judge hires clerks, and the judge's desirability.

Our model shows that the least desirable[46] employers strictly prefer any unraveling equilibrium to an assortative matching that is determined once all pertinent information about

[45] In our model, randomly matching the J best students at T to employers is such a mechanism. A more general case is studied in Halaburda (2010). She first considers an equilibrium with unraveling, and defines a matching mechanism that replicates the outcome of the unraveling equilibrium. This mechanism is incentive compatible but Pareto-inefficient. She can thus adjust it to make all participants in the market weakly better off, while retaining incentive compatibility.

[46] We reiterate that we envisage a market in which all positions offered are more desirable than the outside option.

students has arrived. This explains some judges' opposition towards policy interventions in this market. Whether unraveling increases or decreases the total production of the market depends on whether student skill and employer desirability are complements or substitutes. If a larger number of employers compete for the same set of students, this additional competitive pressure exacerbates unraveling.

Because each equilibrium in our model yields the same equilibrium hiring times and the same expected utility for each player, our model dampens hopes that 'soft' interventions, such as non-binding attempts to coordinate hiring on a particular point in time, could alleviate unraveling. More surprisingly, perhaps, instituting a centralized clearinghouse at the start date of employment will not affect unraveling either, regardless of the assignment mechanism employed.[47] An intervention that *can* affect equilibrium behavior consists of creating a discontinuity in the availability of student information to employers. It is crucial for this to be done boldly; otherwise it may exacerbate unraveling. Implementable policies include withholding existing student information from prospective employers, limiting the production of information about students, and instituting a nationwide test that provides a sudden surge of information about students' abilities at a single point in time.

Behavior in our model contrasts with behavior in two-period models of early offers which dominate the literature. While unraveling serves to thin out the market in our model, competition in two-period models can be avoided partially at best, and hence markets necessarily remain thick. It is the more desirable employers who benefit from the fact that the less desirable employers only have limited opportunities to avoid competition. When we restrict our model to two periods we identify equilibria that differ depending on how much information is available in the earlier period. If there is little information about students, only less desirable employers hire early; if there is a large amount of information, there is no pure strategy equilibrium. While some two-period models find mechanisms that eliminate unraveling, these mechanisms depend on the fact that the informational loss associated with hiring early is discrete. Such a policy is less effective in our model, in which information arrives gradually.

[47]But given that we study a market in which the unique pairwise stable matching would result if all employers postponed hiring to a point in time just before the start of employment, this result is perhaps expected.

Further research might profitably extend the study of unraveling over time by focusing on students. Perhaps the most significant contribution could be made by studying unraveling over time when workers' wages are up for negotiation.

Acknowledgments *We thank Muriel Niederle, Al Roth, and Andy Skrzypacz for their invaluable advice, and Mohammed Akbarpour, Tim Bresnahan, Itay Fainmesser, John Hatfield, Paul J. Healy, Matthew Mitchell, Joshua Mollner, Mikhail Panov, Florian Scheuer, and Pavel Zryumov for helpful comments and suggestions.*

References

Administrative Office of the United States Courts (2013). Key OSCAR Statistics. https://oscar.uscourts.gov/_assets/FY_2013_Key_OSCAR_Statistics.

Alden, W. (2014). A mad scramble for young bankers. *The New York Times*.

Anonymous (03/05/2013). Which judges are the feeder judges? `http://www.top-law-schools.com/forums/viewtopic.php?f=34&t=202674`.

Ante, S. E. (2012). Revenge of the nerds: Tech firms scour college campuses for talent. *Wall Street Journal*.

Avery, C., Jolls, C., Posner, R. A., and Roth, A. E. (2001). The market for federal judicial law clerks. *The University of Chicago Law Review*, 68(3):793–902.

Bonica, A., Chilton, A. S., Goldin, J., Rozema, K., and Sen, M. (2017). The political ideologies of law clerks. *American Law and Economics Review*, 19(1):96–128.

Collins, J. N. (2012). Salaries for new lawyers: An update on where we are and how we got here. *National Association of Law Placement Bulletin*, pages 1–4.

Damiano, E., Hao, L., and Suen, W. (2005). Unraveling of dynamic sorting. *Review of Economic Studies*, 72:1057–1076.

Davis, J. M. (2017). The short and long run impacts of centralized clearinghouses: Evidence from matching teach for america teachers to schools. *unpublished*.

Doval, L. (2017). A theory of stability in dynamic matching markets. *unpublished*.

Du, S. and Livne, Y. (2013). Rigidity of transfers and unraveling in matching markets. *Mimeo*.

Echenique, F. and Pereyra, J. S. (2016). Strategic complementarities and unraveling in matching markets. *Theoretical Economics*, 11.

Fainmesser, I. P. (2013). Social networks and unraveling in labor markets. *Journal of Economic Theory*, 148(2013):64–103.

Ginsburg, T. and Wolf, J. A. (2003). The market for elite law firm associates. *Florida State University Law Review*, 31:909.

Halaburda, H. (2010). Unraveling in two-sided matching markets and similarities of preferences. *Games and Economic Behavior*, 69:365–393.

Haruvy, E., Roth, A. E., and Unver, M. U. (2006). The dynamics of law clerk matching: an experimental and computational investigation of proposals for reform of the market. *Journal of Economic Dynamics and Control*, 30:457–486.

Hautmann, T. (2013). 20-months-old baby becomes world's youngest pro footballer. *Fox Soccer Blog*.

Kadam, S. V. and Kotowski, M. H. (2017). Multi-period matching. *unpublished*.

Kagel, J. H. and Roth, A. E. (2000). The dynamics of reorganization in matching markets. *The Quarterly Journal of Economics*, 115(1):201–235.

Kozinski, A. (1991). Confessions of a bad apple. *The Yale Law Journal*, 100(6):1707–1730.

Kramer, L. (2013). Open Letter to Federal Judges About Clerkships.

Kueppers, C. (2016). A disciplinary association aims to rein in a 'chaotic' hiring calendar. *Chronicle of Higher Education*.

Lat, D. (2011). Clerkship bonus watch: Buckleysandler to $60,000! `http://abovethelaw.com/2011/03/clerkship-bonus-watch-buckleysandler- to-60000/`.

Lat, D. (2013a). The Law Clerk Hiring Plan, R.I.P. `http://abovethelaw.com/2012/06/the-law-clerk-hiring-plan-r-i-p/`.

Lat, D. (2013b). Which law firm won the SCOTUS clerk sweepstakes? `http://abovethelaw.com/tag/clerkship-bonuses/`.

Li, H. and Rosen, S. (1998). Unraveling in matching markets. *The American Economic Review*, 88(3):371–387.

Li, H. and Suen, W. (2000). Risk sharing, sorting, and early contracting. *Journal of Political Economy*, 108(5):1058–1091.

Li, H. and Suen, W. (2004). Self-fulfilling early-contracting rush. *International Economic Review*, 45(1):301–324.

Lindvall, T. (2002). *Lectures on the Coupling Method*. Dover Publications, Mineola, New York.

McDaniel, K., Drummond, D., and Walia, M. (12/08/2013). Clerkship notification blog. `http://cnb09-10.blogspot.com`.

Morse, R. (2014). Grads of these law schools get the most judicial clerkships. *U.S. News & World Report*.

Niederle, M., Proctor, D. D., and Roth, A. E. (2006). What will be needed for the new gastroenterology fellowship match to succeed? *Gastroenterology*, 130:218–224.

Niederle, M., Roth, A. E., and Unver, M. U. (2013). Unraveling results from comparative demand and supply: An experimental investigation. *Games*, 4:243–282.

Okudaira, H. (2017). Regulating the timing of job search: Evidence from new college graduates in japan. *unpublished*.

Online System for Clerkship Application and Review (2013a). Federal law clerk hiring plan information. `https://oscar.uscourts.gov/assets/Federal_Law_Clerk_Hiring_Plan_for_Download.pdf`.

Online System for Clerkship Application and Review (2013b). OSCAR De-links from the Hiring Plan. `https://oscar.uscourts.gov/blog_post/_1/21/OSCAR_De-links_from_the_Hiring_Plan`.

Ostrovsky, M. and Schwarz, M. (2010). Information disclosure and unraveling in matching markets. *American Economic Journal*, 2(2):34–63.

Pan, S. (2017). Exploding offers and unraveling in two-sided matching markets. *International Journal of Game Theory*, pages 1–23.

Popper, N. (2014). Commiting to play for a college, then starting 9th grade. *The New York Times*.

Roth, A. E. (1984). The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92(6):991–1016.

Roth, A. E. and Peranson, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, 89(4):748–80.

Roth, A. E. and Xing, X. (1994). Jumping the gun: Imperfections and institutions related to the timing of market transactions. *The American Economic Review*, 84(4):992–1044.

Strauss, D. M. (2015). Top ten tips to the judicial clerkship interview. http://abovethelaw.com/career-files/ten-tips-to-the-judicial-clerkship-interview/.

Suen, W. (2000). A competitive theory of equilibrium and disequilibrium unraveling in a two-sided matching. *RAND Journal of Economics*, 31(1):101–120.

Tobias, C. (2013). Tips for capturing 2014 federal court clerkships. *Vanderbilt Law Review En Banc*, 66:145–152.

Vidaurre, J. and Campbell, J. (2017). The case for an epilepsy and clinical neurophysiology match. *Pediatric Neurology*, 72:5–6.

Wald, P. M. (1990). Selecting law clerks. *Michigan Law Review*, 89(1):152–163.

Wetz, R. V., Seelig, C. B., Khoueiry, G., and Weiserbs, K. F. (2010). Out-of-match residency offers: the possible extent and implications of prematching in graduate medical education. *Journal of Graduate Medical Education*, 2(3):327–333.

Wikipedia (2013). List of law clerks of the supreme court of the US. `http://en.wikipedia.org/wiki/List_of_law_clerks_of_the_Supreme_Court_of_the_United_States`.

www.top-law-schools.com (2013-2015). Clerkship application thread. `http://www.top-law-schools.com/forums/viewtopic.php?f=34&t=202465` and `http://www.top-law-schools.com/forums/viewtopic.php?t=214750` and `http://www.top-law-schools.com/forums/viewtopic.php?f=34&t=248994`.

Yale Law School Career Development Office (2008). Judicial clerkships in the U.S. `http://www.law.yale.edu/documents/pdf/CDO_Public/cdo-JCpublic2008.pdf`.

Yu, F. (2017). HEP postdoc rumor mill. `https://sites.google.com/site/postdocrumor/`.

APPENDIX
FOR ONLINE PUBLICATION

A Extensions

Our model makes two simplifying assumptions. First, students are not allowed to reject offers. Second, employers only know the properties of the stochastic process according to which student skill levels evolve, but do not observe the current realization of the student skills. Sections A.1 and A.2 relax each of these assumptions in turn. Section A.3 then shows that our main results regarding the dispersion of hiring times and the order in which employers exit can also be obtained as the limit of a sequence of discrete time games. For tractability, we consider the case of two employers.

A.1 Allowing for Student Rejections

Here we show that when students can reject offers, (i) a pure strategy subgame perfect equilibrium always exists, (ii) unraveling may be more or less severe when students can reject offers relative to when they cannot, (iii) as students' utility from the outside option decreases (while their valuations of employers remain constant), the model with student rejections converges to the model without student rejections. This last property illustrates that the model in the main text is one of *elite* entry level labor markets, in which each position is preferred to the outside option.

Assumptions All students assign utility u_1 to a match with employer 1, u_2 to a match with employer 2, and u_0 to being unemployed, with $u_1 \geq u_2 \geq u_0$. Without loss of generality, we set $u_1 = 1$ and $u_2 = 0$. Students have the same information about the evolution of student skills as employers. That is, the process according to which student skills evolve is common knowledge, but students do not possess any information about the current realization of students' skills, not even their own. Once a student receives an offer, however, he can deduce, that he currently is the most highly skilled available student. We use $p^i(t, t')$ to denote the probability that the student who is ranked first at time t will be ranked i at time $t' > t$. It is clear that $p^i(t, t')$ is continuous for all i, and first-order stochastically increasing in t for any fixed t'.[48]

An employer can make an offer to the best available student by playing an action $e \in E$, as in section 3. Whenever an employer j does so, time is halted, the best available student learns the identity of the employer making the offer, and must decide immediately whether

[48] That is, $p^1(t, t')$ and $p^1(t, t') + p^2(t, t')$ are both increasing in t.

or not to accept this employer's offer. If he accepts, he and employer j are matched, and exit the game. Otherwise both players remain in the game. Hence, offers are exploding. We assume that in case of indifference, students accept an offer. Students do not observe offers that are made to other students if they are rejected. All other past actions of all agents are common knowledge.

Note that accepting an offer from employer 1 is a weakly dominant strategy for any student, and so is accepting the best offer received at time T. Hence, we can represent a student's strategy as a mapping $\mathcal{H} \to \{a, r\}$ where a and r are the actions of accepting and rejecting an offer by employer 2, respectively. Additionally, an employer's strategy of making an offer to the currently second best available student at any time $t < T$ is strictly dominated by the strategy of hiring the best available student at T. Therefore, we lose no generality by assuming that employer's actions at each point in time t are whether or not to make an offer to the currently best available student. For tractability we only consider equilibria that are symmetric in the sense that all students play the same strategy. For notational convenience, instead of writing "the currently best available student", we simply write "the student".

Analysis Suppose that the student knows that the next point in time at which employer 1 will make an offer to the then best available student is t. We can then derive the last point in time $\tau(t)$ at which the student would accept an offer from employer 2. From accepting employer 2's offer, a student receives utility u_2. Comparing this to the expected utility from rejecting the offer, we obtain that

$$\tau(t) \quad = \quad \max\{t' : p^1(t', t)u_1 + p^2(t', t)u_2 + (1 - p^1(t', t) - p^2(t', t))u_0 \leq u_2\}$$

We characterize the equilibria of this game depending on the value of $\tau(T)$. As in the case without student rejections, we define t_1 by $x_1(t_1) = x_2(T|t_1)$. Additionally, we define $\bar{t}$ by $x_1(\bar{t}) = x_2(T|T)$ (see figure 5). We distinguish between three cases, depending on the location of $\tau(T)$ relative to t_1 and $\bar{t}$.

Case 1 The case $\tau(t) < \bar{t}$ obtains if u_0 is high, so that students have a strong preference for employer 1 over employer 2 relative to their fear of unemployment. Hiring before $\bar{t}$ is not attractive for either employer: postponing hiring until T yields a higher expected payoff even for the employer who will match to the second best student at T. Employer 2's offer to the best student will be rejected if made at any time $t \in (\tau(T), T)$. Hence employer 1 faces no competition in this time interval, and thus has no reason to make an offer. Thus, the student's assumption that no offer will be made in the interval $(\tau(T), T)$ is correct. Hence, there is an equilibrium without unraveling in which employers will assortatively match to students at T.

Case 2 Next, we consider the case $\bar{t} \leq \tau(T) \leq t_1$, which obtains if u_0 is in an intermediate range such that students have somewhat weaker preferences for employer 1 over employer 2 relative to their fear of unemployment. Employer 2 prefers to make an offer to a student as late as possible. But he also strives to avoid competition by employer 1. Hence, employer 2 will find it optimal to make an offer to the best student at $\tau(T)$, and this student will accept. Employer 1 faces no competition in the time interval $(\tau(T), T)$, and hence will not make any offer during this interval. Moreover, making an offer at or before $\tau(T)$ is a weakly dominated strategy for employer 1. This justifies the student's assumption that no offer will be made in the interval $(\tau(T), T)$. Notably, in this equilibrium, the students' ability to reject offers exacerbates unraveling relative to the case without student rejections.

Case 3 Finally, we study the case $t_1 < \tau(T)$, which is most similar to the case without student rejections. A pure strategy subgame perfect equilibrium always exists, since giving the students the opportunity to reject offers essentially renders the model into a discrete time game, through the following argument. By assumption, at time $\tau(T)$, the last point in time at which the student would accept employer 2's offer, exiting beats staying in the game for both employers. Thus employer 1 will make an offer to the student at $\tau(T)$. Consequently, the student will not accept an offer from employer 2 sufficiently shortly before $\tau(T)$, since he knows that he will receive an offer from employer 1 at time $\tau(T)$. There will, however, be a latest time before $\tau(T)$, denoted $\tau^2(T)$, at which the student would accept employer 2's offer. If that point in time is is again strictly later than t_1^*, exiting at $\tau^2(T)$ beats waiting for both employers, and hence, employer 1 will make an offer to the student at that point in time, and we can iterate the previous argument. Thus, the students' strategy essentially renders the game into a discrete time game between employers, since employer 2 can only possibly exit at times $\tau^n(T)$. In equilibrium, the first exit occurs at one of the two grid points that are closest to t_1. Which one it is depends on whether exiting at the grid point just before t_1^* is better or worse for employer 2 than being preempted by employer 1 at the first grid point after t_1^*. As always, the second exit occurs at T. Moreover, the discrete time grid that emerges becomes increasingly fine as the utility u_0 from the outside option decreases, since the student becomes increasingly unwilling to wait for a any given amount of time in hope for an offer from employer 1. Hence, as u_0 approaches $-\infty$, the time of first exit approaches t_1. The order of exit in the limit is indeterminate in the following sense: For each $\underline{u} \in \mathbb{R}$ a value $u_0' < \underline{u}$ can be found such that employer 1 is first to exit, and another value $u_0'' < \underline{u}$ can be found such that employer 2 is first to exit.

To formally describe the equilibrium in this case, we set $\tau^0(T) = T$. Inductively, we then define $\tau^n(T) = \tau(\tau^{n-1}(T))$ as the last point in time strictly before $\tau(\tau^{n-1}(T))$ at which the student is willing to accept employer 2's offer if he knows that he will receive an offer from employer 1 at time $\tau^{n-1}(T)$ if he is still the best available student at that time. We call the points in time $\{\tau^n(T) | n \in \mathbb{N}\}$ *threat times*. Define $\tau_+^* = \min\{\tau^n(T) : \tau^n(T) \geq t_1, n \in \mathbb{N}\}$, and $\tau_-^* = \tau(\tau_+^*)$

We define the players' strategies as follows:

- The student accepts an offer from employer 2 at all threat times $\tau^n(T) \geq \tau_-^*$, and rejects it otherwise.
- Employer 1 plays e_1 at each threat time $\tau^n(T) \geq \tau_+^*$, and plays s otherwise.
- Employer 2 plays e_1 at each threat time $\tau^n(T) \geq \tau_+^*$, additionally plays e_1 at threat time τ_-^* if $x_1(\tau_-^*) \geq x_2(T|\tau_+^*)$, and plays s otherwise.

We show that this pure strategy profile is subgame perfect Nash equilibrium. Our argument relies on the fact that for any $t > 0$ the interval $[t, T]$ contains a finite number of threat times. To see this, suppose that $\tilde{t} > 0$ is an accumulation point of threat times. This means that for each $\epsilon > 0$ we have

$$p^1(\tilde{t}, \tilde{t} + \epsilon)u_1 + p^2(\tilde{t}, \hat{t} + \epsilon)u_2 + \big(1 - p^1(\tilde{t}, \tilde{t} + \epsilon) - p^2(\tilde{t}, \tilde{t} + \epsilon)\big)u_0 > u_2$$

Note that $p_1(\tilde{t}, \tilde{t}) = 1$. Therefore, the left hand side of the above expression equals u_1 when $\epsilon = 0$, which strictly exceeds u_2. This contradicts the fact that $(p_1(t', t), p_2(t', t))$ is continuous in (t', t).

We now show that no player has an incentive to unilaterally deviate. Clearly, making an offer at any time $t \in (\tau^1(T), T)$ is suboptimal for employer 1, since at any such point in time an offer by employer 2 would be rejected, so that this employer is no serious competitor at such a time. If $\tau^1(T) > t_1$, then employer 2 will make an offer at that time. Because in this case, $x_1(\tau^1(T)) > x_2(T|\tau^1(T))$, employer 1 will also make an offer at that time. Therefore, for any point in time $t \in (\tau^2(t), \tau^1(T))$, rejecting employer 2's offer is rational for the student, and hence employer 1 will not make an offer at any such point in time. But since the student will accept employer 2's offer if he receives it at time $\tau^2(T)$, as long as $x_1(\tau^2(T)) > x_2(T|\tau^2(T))$, employer 1 will also make an offer at that point in time. When the first offer will be made, an by whom it will be made now depends on the exact location of τ_+^* and τ_-^*. By definition of τ_+^*, if both employers are still in the game at that point in time, both employers will play e_1, employer 1 will exit, and employer 2 will earn expected payoff $x_2(T|\tau_+^*)$. Instead, by preempting employer 1 and playing e_1 at τ_-^*, he can instead obtain expected payoff $x_1(\tau_-^*)$.

In conclusion, depending on the precise value of u_0, one of the following two outcomes is consistent with a subgame perfect pure strategy equilibrium profile:

- Employer 2 plays e_1 at τ_-^*, the student accepts, and employer 1 exits at T.
- Both employer 1 and employer 2 play e_1 at τ_+^*, the student employer 1's offer, and employer 2 exits at T.

Finally, it is clear that as u_0 approaches $-\infty$, the distance between any two threat times approaches 0. In particular, $|\tau_-^* - t_1|$ and $|\tau_+^* - t_1|$ approach 0. In this sense, the model with student rejections converges to the model without student rejections.

A.2 Fully Observable Student Skill Levels

Our model is tractable because employers cannot condition their strategies on the current realization of students' skills – they possess no information that would enable them to do so. In this section we relax this assumption. Employers have full knowledge of the current realization of all students' skill levels at each point in time. We formally derive a subgame perfect equilibrium in the case of two employers. We find that hirings are still dispersed in time, and any employer faces the same ex ante expected payoff from hiring at any of these points in time in equilibrium. By arguments that parallel those in sections 6.2 and

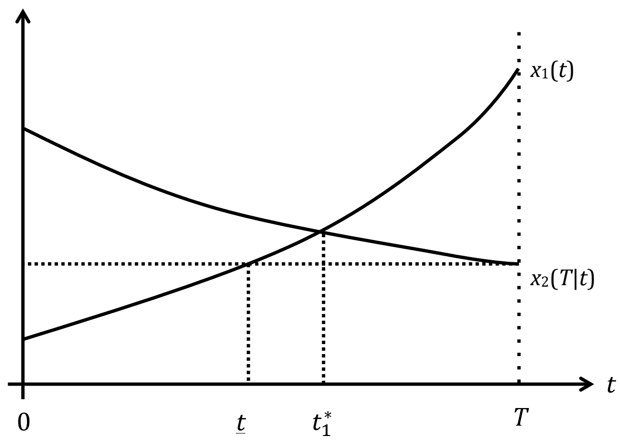

Figure 5: The model with two employers when students can reject offers. The first exit never occurs before $\underline{t}$, but may occur between $\underline{t}$ and t_1^*.

6.1, continuous changes in the speed of arrival of information have no effect on equilibrium payoffs, and the mechanism that is employed by a centralized clearinghouse has no effect on equilibrium behavior.

Setup Both employers can, at any time, make an offer to any student i of their choosing by playing action e_i, or they can choose to play action s and remain in the game. Students must accept any offer they receive, except in the case that they receive multiple offers. In this case, the more attractive employer 1 is assigned the student. For simplicity, we assume that an employer's payoff from failing to hire a student is $-\infty$.

Analysis Let $x_1(t, q)$ denote the skill level that an employer can expect at t if he hires the best student at t, given that the vector of student skill level at that time is q. By the martingale property of the Brownian motion it is immediate that

$$x_1(t, q) = \max_i q_i$$

Once a student has been hired, the best response of the remaining employer is to wait until T and hire the best available student then. Let t be the point in time at which the first student was hired. Let

$$x_2(t,q) = E_t\Big[\max\big\{q_i(T) : 1 \leq i \leq I, i \neq \overline{\text{argmax}}_i q_i(t)\big\}\Big]$$

Here, the function $\overline{\text{argmax}}$ returns the lowest-index maximizer of its argument. $x_2(t,q)$ is the expected value of this strategy conditional on the information available at t. It depends on the vector of skill levels q that prevailed at t, and on the distance between t and the start date of employment, T.

As in our baseline model, we would like to define a point in time at which exiting the game yields the same expected payoff as remaining in the game. This will now depend on the realizations of the student skill levels. Specifically, we define $A \subset [0,T] \times \mathrm{R}^I$ as the set of pairs (t,q) such that the expected value from hiring the best student is equal to the expected value from the continuation strategy $x_2(t,q)$,

$$A = \big\{(t,q) \in [0,T] \times \mathrm{R}^I : x_1(t,q) - x_2(t,q) \geq 0)\big\}$$

The set A characterizes a subgame perfect equilibrium of the game.[49]

Proposition 4. *Define employers' strategies σ_1 and σ_2 as follows. If no student has yet been hired and $(t,q) \in A$, or if a student has been hired and $t = T$, then $\sigma_j(t,q) = e_{\overline{\text{argmax}}_i q_i}$. Otherwise, $\sigma_j(t,q) = s$. Then, $\sigma = (\sigma_1, \sigma_2)$ is a subgame perfect equilibrium.*

Proof. Suppose employer $-j$ plays strategy σ_{-j}. It is immediate that σ_j is a best response to σ_{-j} if and only if σ_j prescribes waiting with hiring until T if employer $-j$ has already exited. Moreover, by the definition of A, employer j strictly prefers playing s to playing e_1 at any $(t,q) \notin A$, and weakly prefers playing e_1 at any $(t,q) \in A$. Consequently, σ_j is a best response to σ_{-j}. □

Observe that each employer's ex ante expected in this equilibrium is the same, and that the points in time at which students are hired are dispersed in time, with the last student being hired at T. This parallels parts (i) and (ii) of corollary 1 in the main text. Moreover, it is apparent that continuous changes in the speed of arrival of information have no effect

[49] Observe that the boundary of A is a one-dimensional smooth submanifold of $[0,T] \times \mathrm{R}^I$. By setting x_3 as the conditional expectation of the maximal student skill at the point in time at which the process hits A, we could, in principle, generalize this analysis to 3 employers, and use induction to extend the analysis to n employers. Doing so requires one to prove a theorem about the smoothness of hitting distributions of I-dimensional Brownian motions on submanifolds of $[0,T] \times \mathrm{R}^I$.

on equilibrium payoffs. It is also apparent that at most one student will be hired at any given point in time, so that the mechanism that is employed by a centralized clearinghouse will have no effect on equilibrium behavior.

A.3 Discrete Time Limits

Here we show that our main results – the dispersion of hiring times, and the indeterminate order in which employers exit the market – are not an artifact of the continuous time formulation we adopt. To do so, we derive the same results as the limit of discrete time games.

Let $(\mathcal{T}^n)_n$ be a sequence of partitions of $[0,T]$ into intervals $[\tau_k^n, \tau_{k+1}^n)$ and the element $\{T\}$, such that l_n, defined as the length of the widest interval in element $\mathcal{T}^n$, converges to 0 as n approaches infinity. For each n define $\tilde{\tau}_+^n = \min\{\tau_k^n : \tau_k^n \geq t_1\}$ and $\tilde{\tau}_-^n = \max\{\tau_k^n : \tau_k^n < t_1\}$.

We derive the set of subgame perfect equilibria for each n. If one employer has exited previously, the remaining employer finds it optimal to stay in the game until T. Hence, we can summarize each employer's strategy by noting the first point in time at which this employer exits if both employers are still in the game. If the more desirable employer 1 would exit at t_k^n, then, as long as $x_1(t_{k-1}^n) > x_2(T|t_k^n)$, the less desirable employer 2 finds it optimal to preempt employer 1 by exiting at t_{k-1}^n. Similarly, if the less desirable employer 2 finds it optimal to exit at t_{k-1}^n, then, as long as $x_1(t_{k-1}^n) > x_2(T|t_{k-1}^n)$, the more desirable employer 1 finds it optimal to preempt employer 2 by exiting at t_{k-1}^n too. For all n with $\tilde{\tau}_+^n > t_1$, backwards induction thus implies that in equilibrium either (i) employer 2 exits at $\tilde{\tau}_-^n$ with employer 1 exiting at T, or (ii) that employer 1 exits at $\tilde{\tau}_+^n$ with employer 2 exiting at T. Case (i) applies if $x_1(\tilde{\tau}_-^n) > x_2(T|\tilde{\tau}_+^n)$. This, in turn, holds if t_1 is located at the lower end in the interval $[\tilde{\tau}_-^n, \tilde{\tau}_+^n)$. Case (ii) applies if the opposite inequality holds, which holds if t_1 is located at the upper end in the interval $[\tilde{\tau}_-^n, \tilde{\tau}_+^n)$. (If $\tilde{\tau}_+^n = t_1$, then there are two equilibria; one employer exits at t_1, and the other exits at T.)

Clearly, both $\tilde{\tau}_-^n$ and $\tilde{\tau}_+^n$ converge to t_1, and in each subgame equilibrium one employer exits at T. This shows that in the limit of the discrete time games the hiring times are given as in the continuous formulation. Additionally, continuity of the functions $x_1(\cdot)$ and $x_2(T|\cdot)$ allows us to choose the sequence $(\mathcal{T}^n)_n$ both such that $x_1(\tilde{\tau}_-^n) > x_2(T|\tilde{\tau}_+^n)$ holds for each n, or that the opposite inequality holds for each n. Thus, both orders of employers exiting the market can be sustained as a limit of discrete time games.

Finally, note that the difference in the expected productivity of the students that employers 1 and 2 are matched to approaches zero from below, and does so monotonically if the sequence of partitions $(\mathcal{T}^n)_n$ is nested. Hence, the more frequent the opportunities for employers to hire students, the smaller the difference in expected worker productivity across employers.

B Numerical Simulation of the Model

To assess the extent of unraveling that our model predicts, and the welfare loss associated with unraveling, we simulate our model numerically. We set $T = 1$ and interpret the equilibrium hiring times as the fraction of time to the start of employment T that has passed.

For welfare analysis we interpret our results in terms of information used or lost in the following sense. We assume that employer desirability and student skills are neither substitutes nor complements: $\alpha_j = 1$ for all j. The distribution of a randomly selected student's skill level at $T = 1$ is standard normal, and hence has standard deviation of 1. Hence, if employer j obtains payoff π, this means that the skill level of the student to which he is assigned, on average, exceeds the expected skill level of a randomly selected student by π standard deviations.

We calculate the conditional expectation functions x_j using Monte Carlo simulations with 50,000 independent realizations of the stochastic processes.[50] We simulate our model with 2, 3 or 4 employers and with 10, 20 or 50 students.[51]

Figure 6 presents the vector of equilibrium hiring times for each market. The extent of unraveling is substantial. In each of the markets considered, the first student is hired before 80% of the time has passed, and when the ratio of students to employers is low, the first student is hired before 30% of the time has passed. The dispersion of hiring times is

[50] We calculate the equilibrium hiring times on a uniform grid with 50 elements.

[51] The computational burden increases exponentially with the number of employers, since each employer increases the dimensionality of the problem. Hence, we consider markets with a small number of employers.

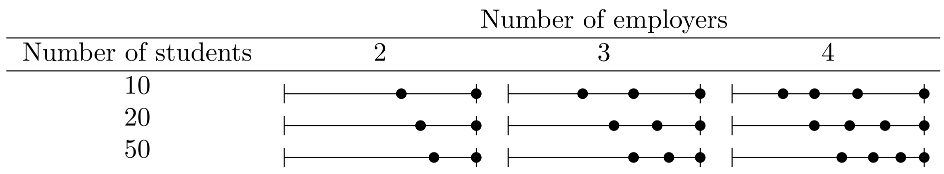

Figure 6: Equilibrium hiring times. Each horizontal bar represents the time interval $[0, T]$. Black dots represent equilibrium hiring times.

substantial and hiring times do not bunch. Moreover, the first hiring time decreases with the number of employers and increases with the number of students.[52]

It would seem that such early hiring times are associated with substantial welfare losses. Our simulations, however, suggest otherwise. Figure 7 reports, for each market, the expected production achieved by the average employer. Because $\alpha_j = 1$ for all j, this equals the expected student skill level obtained. This is contrasted to the employer's expected payoff from an assortative matching of the J best students to employers at T. To interpret the numbers, note that 0 is the expected student skill level that the average employer would obtain from a random assignment of students to employers. Relative to this benchmark, the production attained in equilibrium exceeds 87% of the production that could be obtained by an assortative matching at T (the maximal possible production) in each of the markets analyzed. Moreover, the percentage of maximal production attained decreases with the intensity of competition: it decreases with the number of employers for a given number of students, and it increases with the number of students for a given number of employers.[53]

The unraveling equilibrium is associated with substantial redistribution. In the case of 4 employers and 10 students, from an assortative matching at T the most desirable employer expects almost double the student skill level he expects in equilibrium. In contrast, the least desirable employer expects just about half the expected student skill level in an assortative matching relative to equilibrium. Moreover, the smaller the number of students relative to the number of employers, the larger is the extent of redistribution.

[52] We also observe that the first hiring time is not simply a function of the ratio of employers to students: with 2 employers and 10 students, the first hiring time is 0.61 whereas with 4 employers and 20 students it is 0.43.

[53] We can also infer the reason for the latter effect from our model: the absolute loss in mean productivity due to unraveling varies very little with the number of students, but the mean production that can be achieved from a matching at T increases substantially in the number of students.

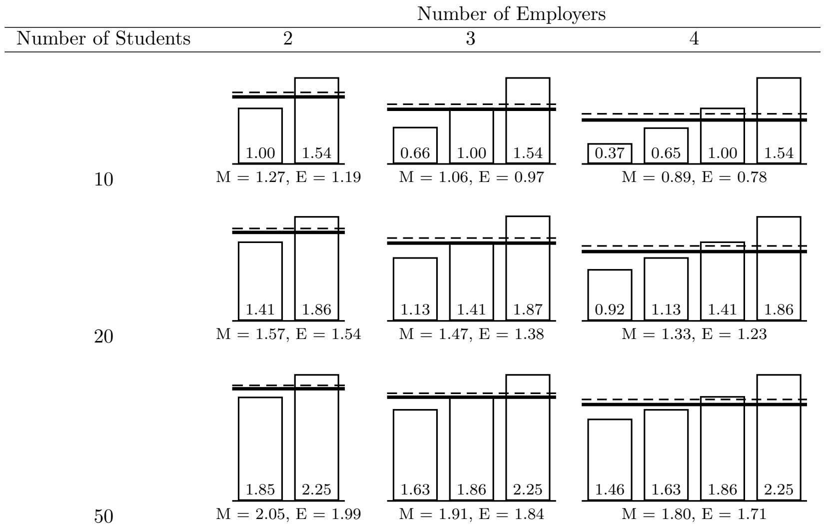

Figure 7: Employer payoffs. The thick vertical line depicts the payoff in the unraveling equilibrium. The dashed line depicts the average payoff form an assortative matching at T. The bars depict the payoff of each employer from an assortative matching at T depending on his desirability.

C Data Sourcing and Encoding

Data sourcing. While we only use data for positions starting in 2009, the Clerkship Notification Blog also features partial data about positions starting in 2010. The data for 2010, however, contain a considerably smaller number of entries, likely due to competing blogs. Unfortunately, data from the blog that ultimately succeeded the Clerkship Notification Blog ("Law Clerk Addict") is no longer available.

We considered an additional data source, and a further proxy for the desirability of the judges. First, the internet blog "Top Law Schools" (Anonymous, 2013) provides a partial account of the number of clerks sent to the Supreme Court by individual circuit judges between 2009 and 2013. As an alternative proxy we considered the number of times a judge was mentioned in the Clerkship Notification Blog. We expect this to be correlated with judges' desirability if clerkships with more highly desirable judges are discussed more often. Given that other factors, such as the predictability of a judge's behavior, may also affect how often a judge's name is mentioned, we expect this proxy to be noisy. As with our main proxy, for neither of these alternative measures of desirability do we find a relationship with judges' hiring times.

Data encoding. We code the timing data as follows. For each of the twelve regional federal court of appeals circuits[54] we made note of the exact time and date of every blog entry that included information on hiring and interviewing. We categorized these entries in two dimensions. First, we noted whether claims were being made about (i) students still in law school, (ii) individuals who had previously graduated from law school, or (iii) whether it was unclear. Second, we classified entries as claims that a judge (a) was reviewing applications, (b) was making telephone calls or sending emails to schedule interview times, (c) was interviewing, (d) had made a hire, or (e) had finished the hiring process. We adhered to the following four rules while categorizing our data. 1. Unless blog entries were subsequently claimed to be false or corrected, we assumed every blog entry that was phrased as a statement rather than a question to be evidence that a hiring event occurred. 2. Clear confirmations of events were treated as repeated data points and were not included in our analysis. (For example, while we recorded the entry "Bea called to set up interviews" on September 8, 2008 at 8:48 pm, we did not record the subsequent entry "I can confirm that Bea definitely set up at least two interviews today." on September 8, 2008 at 11:35 pm. Whereas, both of the following entries were treated as unique data points: "Gorsuch has hired two for the current term..." on July 7, 2008 at 8:24 am and "Gorsuch has hired a ... clerk from Harvard." on July 17, 2008 at 3:46 pm.) 3. We did not include very late entries which occurred considerably after the majority of activity had ceased since this was most likely relevant to the 2010-2011 season. 4. We recorded the date of each hire, interview, review, etc. as the time at which the event was recorded on the blog. We did not take into account any comments about when the event occurred (such as "Interviewed with Briscoe last week..."). One exception to this rule is that there were a number of entries that explicitly stated that calls were made at, or a few minutes after, 12:00 pm on September 8th. These blog entries were mostly all posted within a few hours of 12:00 pm and were recorded as 12:00 pm rather than the precise time at which they were posted.

[54] The number of blog entries for the federal circuit court of appeals is very limited, most likely because many federal circuit judges do not hire current law students. For this reason, we exclude these entries from our analysis.

Some data points may concern the hiring of students who have already graduated, and to whom the Hiring Plan does not apply. For the vast majority of the blog entries in our data it was not possible to determine whether they were referring to a current or graduated law student. The fact that the hiring plan was officially abandoned in November 2013 because fewer and fewer judges adhered to it makes us confident that a significant number of our data points refer to current law students rather than graduates.

D Mathematical Appendix

D.1 Definition of mixed stratgies

A (mixed) strategy of an employer j is a function $\sigma_j : \mathcal{H} \to E \cup \{s\}$ that maps a history H_t into a cumulative probability distribution $F^j(\cdot; H_t) \in [0,1]^I$ with support contained in $[t, T]$. The ith component of $F^j(t'; H_t)$ is the probability that employer j plays one of the actions $e_1, \ldots, e_i$ in the interval $[t, t']$ if no other employers have taken any action $e \in E$ in the interval $[t, t')$. Because F^j is a cumulative distribution function, it is right-continuous.[55] This implies that for any pure strategy an employer might have, in any subgame there is a well-defined first point in time at which an employer takes an action $e \in E$.[56]

We require each strategy σ_j to be internally consistent in the following sense: fix a history H_t, and define, for all $\tilde{t} \geq t$, the history $\tilde{H}_{\tilde{t}}$ as the history which coincides with H_t up to and including time t, and in which every employer plays action s at each point in time $[t, \tilde{t}]$. Then for any $t' \in [\tilde{t}, T]$, the probability that $\sigma_j(\tilde{H}_{\tilde{t}})$ assigns to an action $e \in E$ in time interval $[\tilde{t}, t']$ coincides with the probability that $\sigma_j(H_t)$ assigns to the same action e in the same time interval $[\tilde{t}, t']$.

D.2 Proof of proposition 1

We start by defining the notation. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$ be a filtered probability space. Let q be an I-dimensional Brownian motion defined on this space with $q(0) = 0$. For path $\omega \in \Omega$ and time $t \in [0, T]$, we write $q^\omega(t) = (q_1^\omega(t), \ldots, q_I^\omega(t))$, and use W to denote the Wiener measure. $\mathcal{N}(\mu, \sigma^2)$ denotes the unidimensional normal distribution with mean μ and variance σ^2. The symbol $\succsim$ denotes first order stochastic dominance, and $\simeq$ denotes equality in distribution. In the Euclidian space $\mathbb{R}^n$, $\geq$ denotes the product order[57] and $\mathbb{B}_\epsilon(x)$ denotes the ϵ-ball with radius ϵ and center x. For $x \in \mathbb{R}^I$, we let $\overline{\text{argmax}}(x)$ denote the minimal i such that $x_i \geq x_{i'}$ for all $i, i' \in \{1, \ldots, I\}$. We denote the set of available students after previous students have been picked at times $\tau_1, \ldots, \tau_k$ by $a_{k+1}(\tau_1, \ldots, \tau_k)$.

[55] A function $f : \mathbb{R} \to \mathbb{R}$ is *right continuous* at $x \in \mathbb{R}$ if for all sequences $(x_n)_{n \in \mathbb{N}}$ with $x_n > x$ for all $n \in \mathbb{N}$, $\lim_{x_n \to x} f(x_n) = f(x)$.

[56] For instance, a plan of action in which an employer plays s for all $t \leq \hat{t}$ for some $\hat{t}$, and plays e_1 for all $t > \hat{t}$, is not an admissible strategy.

[57] I.e. for $x, y \in \mathbb{R}^n$ we have $x \geq y \Leftrightarrow x_i \geq y_i \; \forall 1 \leq i \leq n$.

We recursively define $a_{k+1}(\tau_1, \ldots, \tau_k)$ by setting $a_1 = \{1, \ldots, I\}$, and

$$a_{k+1}(\tau_1, \ldots, \tau_k) = a_k(\tau_1, \ldots, \tau_{k-1}) \setminus \overline{\text{argmax}}_{i \in a_k(\tau_1, \ldots, \tau_{k-1})} q_i(\tau_k)$$

Additionally, we define

$$i^*(t, a_{k+1}(\tau_1, \ldots, \tau_k)) = \overline{\text{argmax}}_{i \in a_k(\tau_1, \ldots, \tau_k)} q_i(t).$$

Note that both a and i^* are stochastic processes.

The proposition to be proven here concerns properties of the conditional expectation functions. The conditional expectation function $x_{k+1}(t|\tau_1, \ldots, \tau_k)$ is given by

$$\begin{aligned} x_{k+1}(t|\tau_1, \ldots, \tau_k) &= E\left[q_{i^*(t, a_{k+1}(\tau))}(T)\right] \\ &= E\left[E\left[q_{i^*(t, a_{k+1}(\tau))}(T) | \mathcal{F}_t\right]\right] \\ &= E\left[E\left[\max_{i \in a_{k+1}(\tau)} \{q_i(t)\} \middle| \mathcal{F}_t\right]\right] \\ &= E\left[\max_{i \in a_{k+1}(\tau)} \{q_i(t)\}\right] \end{aligned}$$

The second and fourth line follow by the law of iterated expectations, the third by the martingale property of the Brownian motion. We now show that the conditional expectation functions satisfy the five properties in proposition 1.

(i) To show that $x_{k+1}(t|\tau_1, \ldots, \tau_k)$ is jointly continuous in all arguments, choose $t \in [0, T]$ and $\tau \in [0, T]^k$, with $t \geq \tau_k \geq \ldots \geq \tau_1$. Fix a sequence $(t^l, \tau^l)_l$ that converges to (t, τ), such that $t^l \geq \tau_k^l \geq \ldots \geq \tau_1^l$ for all l. We need to show that $\lim_{l \to \infty} |x_{k+1}(t^l|\tau^l) - x_{k+1}(t|\tau)| = 0$.

We define

$$X_l = \max_{i \in a(\tau^l)} q_i(t^l) - \max_{i \in a(\tau)} q_i(t)$$

Note that $\left|x_{k+1}(t^l|\tau_1^l, \ldots, \tau_k^l) - x_{k+1}(t|\tau_1, \ldots, \tau_k)\right| = |E(X_l)|$. By Jensen's inequality, we have $|E(X_l)| \leq E(|X_l|)$. It thus suffices to show that $X_l \to 0$ in the $\mathcal{L}_1$-norm. We proceed in two steps. Lemma 2 shows that X_l converges to 0 almost everywhere. Lemma 3 shows that the sequence X_l is uniformly integrable. A direct application of the Lebesgue-Vitali theorem (theorem 4.5.4 in Bogachev (2007)[58]) then yields $X_l \to 0$ in $\mathcal{L}_1$.

Lemma 2. $P(\lim_{l\to\infty} X_l = 0) = 1$

Proof. Let $\Omega' = \{\omega \in \Omega : q(\omega) \text{ is continuous }\}$. It is a standard result in the theory of Brownian motions that $P(\Omega') = 1$. Let $\Omega'' = \{\omega \in \Omega : q_i(t') \neq q_{i'}(t') \ \forall t' \in \{\tau_1, \dots, \tau_k, t\} \ \forall i \neq i'\}$. Because the Brownian motion has Gaussian increments, $P(\Omega'') = 1$, and hence $P(\Omega' \cap \Omega'') = 1$.

Let $\omega \in \Omega' \cap \Omega''$. Because the path $q^\omega(t')$ is continuous in t', we get that for all $t' \in \{\tau_1, \dots, \tau_k, t\}$ there exists $\epsilon(t') > 0$ such that for all $t'' \in \mathbb{B}_{\epsilon(t')}(t')$ we have $\overline{\text{argmax}}_{i' \in a_k(\tau)} q^\omega_{i'}(t') = \overline{\text{argmax}}_{i' \in a_k(\tau)} q^\omega_{i'}(t'')$.

Because $(\tau^l, t) \to (\tau, t)$, we get that for all sufficiently large l, $a_k(\tau) = a_k(\tau^l)$ and $\overline{\text{argmax}}_{i \in a_k(\tau)} q_i(t) = \overline{\text{argmax}}_{i \in a_k(\tau)} q_i(t^l)$. Then it follows by continuity of q^ω that $q_{i^*(t,a(\tau))}(t) = \lim_l q_{i^*(t,a(\tau^l))}(t^l)$ as was to be shown. □

Lemma 3. *The sequence $(X_l)_l$ is uniformly integrable.*

Proof. For each l, let $(Y^l_i)_{i \in 1,\dots,I}$ denote a sequence of vectors of $\mathcal{N}(0, |t - t^l|)$ random variables, and let $(Z^l_i)_{i \in 1,\dots,I}$ denote a sequence of vectors of $\mathcal{N}(0, \min\{t, t^l\})$ random variables such that for each l, and for all $i \neq i'$, Y^l_i and Z^l_i are independent.

Let $\bar{u}_l = \max_{i \in I} |q_i(t^l)|$, and $u_l = \max_{i \in a(\tau^l)} q_i(t^l)$. Likewise, let $\bar{v}_l = \max_{i \in I} |q_i(t)|$, and $v_l = \max_{i \in a(\tau)} q_i(t)$. Trivially, $\bar{u}_l \geq u_l$, and $\bar{v}_l \geq v_l$.

We can bound $|X^l|$ as follows.

$$\begin{aligned} |X_l| &= |u_l - v_l| \\ &\leq |u_l| + |v_l| \\ &\leq \bar{u}_l + \bar{v}_l \\ &= \max_{i \in 1,\dots,I} |Z^l_i + Y^l_i| + \max_{i \in 1,\dots,I} |Z^l_i| \\ &\leq \max_{i \in 1,\dots,I} |Y^l_i| + 2 \max_{i \in 1,\dots,I} |Z^l_i| \qquad (1) \end{aligned}$$

[58]Bogachev, V. I. (2007) *Measure Theory*, Springer, Berlin.

The second and the last line follow by the triangle inequality, the fourth line is due to Jensen's inequality, and the penultimate line follows by the independent Gaussian increments property of the Brownian motion. By the above, it suffices to show that the sequence $\left(\max_{i\in 1,\ldots,I}|Y_i^l| + 2\max_{i\in 1,\ldots,I}|Z_i^l|\right)_l$ is uniformly integrable.

It is a standard result in probability theory that if Z is a $\mathcal{N}(0,\sigma^2)$ variable, then $E(Z|Z\geq c) = \sigma\frac{\phi(c)}{1-\Phi(c)}$, where ϕ and Φ denote the p.d.f. and c.d.f. of the standard normal distribution, respectively. Thus, $E(|Z|\mathbb{1}_{\{|Z|\geq c\}}) = 2E(Z|Z\geq c)P(Z\geq c) = 2\sigma\phi(c)$.

Note that

$$\begin{aligned} E\left(\left(\max_{i\in 1,\ldots I}|Y_i^l|\right)\mathbb{1}_{\{\max_{i\in 1,\ldots I}|Y_i^l|\geq c\}}\right) &\leq \sum_{i\in 1,\ldots I} E\left(|Y_i^l|\mathbb{1}_{\{|Y_i^l|\geq c\}}\right) \\ &= 2I|t-t^l|\phi(c) \end{aligned}$$

Similarly, $E\left(\left(\max_{i\in 1,\ldots,I}|Z_i^l|\right)\mathbb{1}_{\{\max_{i\in 1,\ldots,I}|Z_i^l|\geq c\}}\right) \leq 2I|\max\{t,t^l\}|\phi(c)$

Combining these results with (1), we can derive the following bound

$$\begin{aligned} \int_c^\infty |X_l|dW &\leq 2\int_c^\infty \max_{i\in 1,\ldots I}|Y_i^l|dF_Y + 4\int_c^\infty \max_{i\in 1,\ldots I}|Z_i^l|dF_Z \\ &\leq \left(2I|t-t^l| + 4I\max\{t,t^l\}\right)\phi(c) \end{aligned}$$

where F_Y and F_z denote the distributions of the random vectors Y and Z, respectively. Because $t^l\to t$, both sequences $\left(|t-t^l|\right)_l$ and $\left(\max\{t,t^l\}\right)_l$ are bounded above , and hence, the sequence $\left(2I|t-t^l| + 4I\max\{t,t^l\}\right)_l$ is bounded above by some $C\in\mathbb{R}$. Therefore,

$$\lim_{c\to\infty}\sup_l \int_c^\infty |X_l|dW \leq \lim_{c\to\infty} C\phi(c) = 0$$

as was to be shown.

□

(ii) To show that $x_{k+1}(t|\tau_1,\ldots,\tau_k)$ is strictly increasing in t, fix $\tau \in [0,T]^k$ with $\tau_k \geq \ldots \geq \tau_1$, and let $t' > t \geq \tau_k$. By the law of iterated expectations,

$$\begin{aligned} x_{k+1}(t'|\tau) &= P\left(i^*(t',a(\tau)) = i^*(t,a(\tau))\right)\cdot \\ &\quad E\left[\max\{q_i(t') : i \in a(\tau)\}|i^*(t',a(\tau)) = i^*(t,a(\tau))\right] \\ &\quad + P\left(i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right)\cdot \\ &\quad E\left[\max\{q_i(t') : i \in a(\tau)\}|i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right] \end{aligned}$$

Now, $E\big[\max\{q_i(t') : i \in a(\tau)\}|i^*(t',a(\tau)) = i^*(t,a(\tau))\big] =$ $E\left[\max\{q_i(t) : i \in a(\tau)\}|i^*(t',a(\tau)) = i^*(t,a(\tau))\right]$ by the martingale property and the conditioning. Moreover, $E\big[\max\{q_i(t') : i \in a(\tau)\}|i^*(t',a(\tau))$ $\neq i^*(t,a(\tau))\big] > E\left[\max\{q_i(t) : i \in a(\tau)\}|i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right]$ by the definition of i^*. Finally, $P\left(i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right) > 0$ because the Brownian motion has Gaussian increments and thus, $q(t')|q(t)$ has full support on $\mathbb{R}^I$. Therefore,

$$\begin{aligned} x_{k+1}(t'|\tau) &> P\left(i^*(t',a(\tau)) = i^*(t,a(\tau))\right)\cdot \\ &\quad E\left[\max\{q_i(t) : i \in a(\tau)\}|i^*(t',a(\tau)) = i^*(t,a(\tau))\right] \\ &\quad + P\left(i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right)\cdot \\ &\quad E\left[\max\{q_i(t) : i \in a(\tau)\}|i^*(t',a(\tau)) \neq i^*(t,a(\tau))\right] \\ &= x_{k+1}(t|\tau) \end{aligned}$$

as was to be shown.

(iii) We prove two lemmas that imply the claim. For each $d \in \{1,\ldots,I\}$, we define the following three mappings. $\varphi_d : \mathbb{R}^d \to \mathbb{R}^{d-1}$ is given by $\varphi_d(x) = (x_1,\ldots,x_{i^*(x)-1},x_{i^*(x)+1},\ldots,x_d)$. Hence, $\varphi_d(x)$ equals x with the highest element removed. For any $1 \leq k \leq d$, we define $\pi_d^k : \mathbb{R}^d \to \mathbb{R}^{d-1}$ by $\pi_d^k(x) = (x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_d)$. Finally, $\psi_d : \mathbb{R}^d \to \mathbb{R}^d$ is the mapping that orders a vector by size, largest dimension first (lowest index first in case of ties). Plainly, for any random vector X, the order statistics of X and $\psi_d(X)$ are identical.

Intuitively, the following lemma shows that when employer j' picks before some time t', the pool of remaining students at t' is better (in the f.o.s.d. sense) than if he picks at t'.

Lemma 4. *Let X and Y be independent $\mathbb{R}^d$-valued random variables. Let $M = \overline{\text{argmax}}_{1\leq i\leq d} X_i$. Then,*

$$\psi_d\left[\varphi_d(X) + \pi_d^M(Y)\right] \succsim \psi_d\left[\varphi_d(X+Y)\right] \tag{2}$$

Proof. Let $U = X + Y$. Let R denote the ranking of U_M. I.e. $R = 1$ if U_M is the highest element of U, $R = 2$ if U_M is the second highest element of U, and so on. Then, letting LH and RH denote the left and right hand sides of equation (2), respectively, we get

$$\begin{aligned} LH &= \left[\big(\psi_d(U)\big)_1, \ldots, \big(\psi_d(U)\big)_{R-1}, \big(\psi_d(U)\big)_{R+1}, \ldots, \big(\psi_d(U)\big)_d\right] \\ RH &= \left[\big(\psi_d(U)\big)_2, \ldots, \big(\psi_d(U)\big)_R, \big(\psi_d(U)\big)_{R+1}, \ldots, \big(\psi_d(U)\big)_d\right] \end{aligned}$$

Hence, $LH_i \geq RH_i$ for all $i \leq R-1$, and $LH_i = RH_i$ for all $i \geq R$. Consequently, $P(LH \geq RH) = 1$ so that the claim is a direct application of Strassen's theorem (theorem 2.4 in chapter IV, Lindvall (2002)). □

Intuitively, the following lemma shows that if the best student i is picked from a pool of students, then the better this pool was before the best student was picked (in the f.o.s.d. sense), the better the pool of the remaining students (in the f.o.s.d. sense).

Lemma 5. *Let X, X', and Y be $\mathbb{R}^d$-valued random variables, such that*

(a) $X \succsim X'$

(b) X and X' are both independent of Y.

Then, $\varphi_d(X+Y) \succsim \varphi_d(X'+Y)$.

Proof. By Strassen's theorem, the stochastic dominance relation $X \succsim X'$ is equivalent to the existence of a coupling $\hat{P}$ such that $\hat{P}(X \geq X') = 1$.[59] Define the random vari-

[59] A coupling of two probability measures P and P' on the same measurable space $(E, \mathcal{E})$ is any probability measure $\hat{P}$ on the product measurable space $(E \times E, \mathcal{E} \otimes \mathcal{E})$ (where $\mathcal{E} \otimes \mathcal{E}$ is the smallest σ-field containing $\mathcal{E} \times \mathcal{E}$) whose marginals are P and P'.

able (Y_1, Y_2) on $\mathbb{R}^d \times \mathbb{R}^d$ such $Y_1 \simeq Y$, $Y_2 \simeq Y$, and $\hat{P}(Y_1 = Y_2) = 1$. Then, $\hat{P}(X+Y \geq X'+Y) = 1$, and by monotonicity of φ_d we obtain $\hat{P}\left(\varphi_d(X+Y) \geq \varphi_d(X'+Y)\right) = 1$. By Strassen's theorem this is equivalent to the claim. □

We now combine lemmas 4 and 5 to show that $x_{k+1}(t|t_1, \ldots, t_k)$ is decreasing in t_j for all $j \in \{1, \ldots, k\}$. Fix $t_1 \leq \ldots \leq t_k$, and fix some $1 \leq j \leq k$. Then, if $j > 1$, let $t'_j \in (t_{j-1}, t_j)$. If instead $j = 1$, let $t'_j \in [0, t_j)$. (For $t'_j < t_{j-1}$, the claim follows by repeated application of the current argument.)

For any subset $a \subseteq \{1, \ldots, I\}$, let $\rho^a : \mathbb{R}^I \to \mathbb{R}^{|a|}$ be given by $\rho^a(x) = (x_i)_{i \in a}$. Let X be a random variable on $\mathbb{R}^{I-j}$ with $X \simeq \rho^{a(t_1,\ldots,t_j)}\big(q(t'_j)\big)$. Let Y be an $(I-j)$-dimensional normal variable with mean zero and covariance matrix $(t_j - t'_j)\mathbb{I}_{I-j}$ that is independent of X, where $\mathbb{I}_d$ is the identity matrix in $\mathbb{R}^d$. Because the Brownian motion has independent Gaussian increments, $\rho^{a(t_1,\ldots,t_j)}\big(q(t_j)\big) \simeq X + Y$.

If the j'th student is picked at t_j, the distribution of the order statistics of the remaining students at t_j is given by $\psi_{I-j}\big(\varphi_{I-j+1}(X+Y)\big)$. If the j'th student is picked at t'_j instead, the distribution of the order statistics of the remaining students at t_j is given by $\psi_{I-j}\big(\varphi_{I-j+1}(X) + \pi^M_{I-j+1}(Y)\big)$, where $M = \overline{\text{argmax}}_i X_i$.

By lemma 4 we thus have that

$$\psi_{I-j}\left(\rho^{a(t_1,\ldots,t_{j-1},t'_j)}\big(q(t_j)\big)\right) \succsim \psi_{I-j}\left(\rho^{a(t_1,\ldots,t_{j-1},t_j)}\big(q(t_j)\big)\right)$$

The above shows that the distribution of the quality of the remaining students at t_j is better if the j'th student is picked at t'_j than if he is picked at t_j.[60] By iterated application of lemma 5 and the fact that the Brownian motion has independent Gaussian increments it now follows that this relation survives to time t in spite of the fact that at times $t_{j+1}, \ldots, t_k$, the best available students are picked off the market. Formally, we obtain

$$\psi_{I-j}\left(\rho^{a(t_1,\ldots,t_{j-1},t'_j,t_{j+1},\ldots,t_k)}\big(q(t)\big)\right) \succsim \psi_{I-j}\left(\rho^{a(t_1,\ldots,t_{j-1},t_j,t_{j+1},\ldots,t_k)}\big(q(t)\big)\right)$$

[60] Observe that the mapping ψ_{I-j} is a mere relabeling of the dimensions of the Brownian motion.

Because the maximum is an increasing function, it thus follows from the definition of first order stochastic dominance that

$$E\left[\max_{i\in a(t_1,\ldots,t_{j-1},t'_j,t_{j+1},\ldots,t_k)} q_i(t)\right] \geq E\left[\max_{i\in a(t_1,\ldots,t_{j-1},t_j,t_{j+1},\ldots,t_k)} q_i(t)\right]$$

as was to be shown.

(iv) Let $q_{(i)}(t)$ denote the ith highest dimension of $q(t)$. Because the increments of the Brownian motion have full support on $\mathbb{R}^I$, $P\left(q_i(T) < q_{(I-J)}(T)|q_i(t) = q_{(I)}(t)\right) > 0$ for all $t < T$. Hence, $\sum_j x_j(T|T,\ldots,T) > \sum_j x_j(t_j|t_1,\ldots,t_{j-1})$, as was to be shown.

(v) For any given path of the Brownian motion, $a(\tau_1,\ldots,\tau_k) \subsetneqq a(\tau_1,\ldots,\tau_{k-1})$. Moreover, for all $i \neq i'$ and for all $t > 0$ we have $P(q_i(t) = q_{i'}(t)) = 0$. Hence, $x_{k+1}(t|\tau_1,\ldots,\tau_k) = E\left[\max_{i\in a(\tau_1,\ldots,\tau_k)} q_i(t)\right]$ $< E\left[\max_{i\in a(\tau_1,\ldots,\tau_{k-1})} q_i(t)\right] = x_k(t|\tau_1,\ldots,\tau_{k-1})$, as claimed.

D.3 Proof of lemma 1

For easier readability, we use the following notation. Consider a history H_t in which k employers have already exited at times $s_1,\ldots,s_k$. We then write

$$\phi_{k+1}(s_1,\ldots,s_k;t) = \inf\{t' \geq t|t' \text{ satisfies the properties in definition 1}\}$$

Moreover, we write

$$g_{k+1}(s_1,\ldots,s_{k-1};t) = x_{k+1}(\phi_{k+1}(s_1,\ldots,s_{k-1};t)|s_1,\ldots,s_{k-1},t) \tag{3}$$

Intuitively, $g_{k+1}(s_1,\ldots,s_{k-1};t)$ is the value of remaining in the game when the first $k-1$ employers have exited at times $s_1,\ldots,s_{k-1}$, and the kth employer exits at time t.

We use induction over k, starting from $k = J$, to prove the following lemma

Lemma 6. *For any $k = 1,\ldots,J$ and for any history H_t in which $k-1$ employers have exited at previous times $h_t = (s_1,\ldots,s_{k-1})$,*

(i) there exits a unique next candidate hiring time $\phi_k(s_1,\ldots,s_{k-1};t)$.

(ii) $g_{k+1}(s_1, \ldots, s_{k-1}; t)$ *is weakly decreasing in* t*, and is weakly decreasing in* $s_1, \ldots, s_{k-1}$ *for all* $s_{k-1} \le t$*.*

To obtain the entire vector of future candidate hiring times associated with a given history H_t with previous hiring times $s_1, \ldots, s_{k-1}$ we then define recursively:

$$\begin{aligned} t_k(H_t) &= \phi_k(s_1, \ldots, s_{k-1}; t) \\ t_{k+j}(H_t) &= \phi_{k+j}\big(s_1, \ldots, s_{k-1}, t_k(H_t), \ldots, t_{k+j-1}(H_t)\big) \end{aligned}$$

thus finishing the proof.

To prove lemma 6, we first prove another lemma. Intuitively, in the case of $n = 1$, lemma 7 states the following: Suppose f is a strictly increasing function of t, and g is a weakly decreasing function of t, and both f and g are continuous and cross at some point t'. Now shift both f and g upwards to some $\hat{f}$ and $\hat{g}$ such that $\hat{f}$ is strictly increasing in t, $\hat{g}$ is weakly decreasing in t, and $\hat{f}$ and $\hat{g}$ intersect at t''. Then $\hat{f}(t'') \ge f(t')$.

Lemma 7. *Let* $\underline{t}, \tilde{t}, \bar{t} \in \mathbb{R}$ *with* $\underline{t} \le \tilde{t} \le \bar{t}$*. Let* $\Theta = \{(\theta_1, \ldots, \theta_n) \in [\underline{t}, \tilde{t}]^n | \theta_1 \le \theta_2 \le \ldots \le \theta_n\}$*. Let* $f, g : [\underline{t}, \bar{t}] \times \Theta \to \mathbb{R}$ *be continuous functions. Suppose that both* f *and* g *are weakly decreasing in the second argument,* θ*, that* f *is strictly increasing, and that* g *is weakly decreasing in the first argument. Moreover, suppose that* $f(\bar{t}, \theta) > g(\bar{t}, \theta)$ *for all* θ*. Define* $t^*(\theta)$ *by*

$$f(t^*(\theta), \theta) = g(t^*(\theta), \theta) \tag{4}$$

if such a value exists and weakly exceeds $\tilde{t}$*, and define* $t^*(\theta) = \tilde{t}$ *otherwise. Then*

(i) $f(t^*(\theta), \theta)$ *is weakly decreasing in* θ*,*

(ii) $t^*(\theta)$ *is continuous.*

Proof.

(i) For all $\theta \in \Theta$, let $G(\theta) = \{(x, t) \in \mathbb{R} \times [t, \bar{t}] : x \le \min\{f(t, \theta), g(t, \theta)\}\}$. If $f(t^*(\theta), \theta) = g(t^*(\theta), \theta)$, then because f is increasing and g is decreasing in its first argument, $f(t^*(\theta), \theta) = \max_{(x,t) \in G(\theta)} x$. If $t^*(\theta) = t$, then $f(t', \theta) \ge g(t', \theta)$ for all $t' \ge t$. Therefore, and because g is weakly decreasing in its first argument, we again have

$f(t^*(\theta), \theta) = \max_{(x,t)\in G(\theta)} x$. Because both f and g are uniformly nonincreasing in the second argument, so is $\max_{(x,t)\in G(\theta)} x$, which completes the proof.

(ii) Let S denote the set of all θ such that a solution to (4) exists. $t^*(\theta)$ is continuous on the interior of S since f and g are both continuous, and it is trivially continuous on the complement of S. It remains to show that it is continuous on the boundary of S. For this it is sufficient to show that $t^*(\theta) = \tilde{t}$ for all θ on the boundary of S. To do so, define $\psi(t, \theta) := f(t, \theta) - g(t, \theta)$. Note that ψ is continuous and strictly increasing in its first argument. Set $\mathcal{T} := \psi^{-1}(\{0\})$. Since ψ is continuous, and since $\mathcal{T}$ is the preimage of a closed set, $\mathcal{T}$ is closed. Consequently, S, which is the projection of $\mathcal{T}$ to Θ, is a closed set. Now choose $\tilde{\theta}$ on the boundary of S. Suppose that $t^*(\tilde{\theta}) > \tilde{t}$. Then, due to monotonicity of ψ in the first argument, $\psi(\tilde{t}, \tilde{\theta}) < 0$. Hence, by continuity of ψ, there exists $\delta > 0$ and $\eta > 0$ such that for all θ in an η-neighborhood of $\tilde{\theta}$ we have $\psi(\tilde{t}+\delta, \theta) < 0$. Hence, by continuity of ψ, and because $\psi(\bar{t}, \theta) > 0$, we have that for all θ in the η-neighborhood of $\tilde{\theta}$, a solution to (4) exists. This contradicts the assumption that $\tilde{\theta}$ was chosen on the boundary of S.

□

Proof. (of lemma 6)

Induction anchor.

(i) Let H_t be a history in which $J - 1$ employers have exited at times $s_1, \ldots, s_{J-1}$. Set $\phi_J(s_1, \ldots, s_{J-1}) = T$. Trivially, this is unique and satisfies the properties in definition 1.

(ii) Observe that by the above, $g_J(s_1, \ldots, s_{J-2}; t) = x_J(T|s_1, \ldots, s_{J-2}, t)$, which is jointly continuous and decreasing in all its arguments by proposition 1.

Induction step.

Suppose lemma 6 holds for all histories H'_t in which k employers have exited. (This is the induction assumption.) We show that it then holds for all histories H_t in which $k - 1$ employers have exited.

(i) Choose a history H_t in which $k-1$ employers have exited, and let $s_1, \ldots, s_{k-1}$ denote the points in time at which they exited. We consider two cases.

Case 1

Suppose for all $t' \geq t$ we have
$x_k(t'|s_1, \ldots, s_{k-1}) \neq x_{k+1}(\phi_{k+1}(s_1, \ldots, s_{k-1}, t)|s_1, \ldots, s_{k-1}, t)$. Part (iii) of definition 1 implies $\phi_k = t$. This shows uniqueness. For existence, it remains to show that $\phi_k = t$ satisfies the remaining properties of definition 1. (i) does not apply. If (iv) applies, then property (ii) follows directly by part (v) of proposition 1. If (iv) does not apply, then we need to show that (ii) is satisfied. Suppose this is not the case. Hence, $x_k(\phi_k|s_1, \ldots, s_{k-1}) < x_{k+1}(\phi_{k+1}(s_1, \ldots, s_{k-1}, \phi_k)|s_1, \ldots, s_{k-1}, \phi_k)$. By part (i) of proposition 1, both sides of this inequality are continuous. By part (v) of proposition 1, $x_k(T|s_1, \ldots, s_{k-1}) > x_{k+1}(T(s_1, \ldots, s_{k-1}, T)|s_1, \ldots, s_{k-1}, T)$. Hence, by the intermediate value theorem of real analysis, there exists $t' \in (t, T)$ such that $x_k(t'|s_1, \ldots, s_{k-1}) = x_{k+1}(t'|s_1, \ldots, s_{k-1}, t')$, in contradiction to the assumption of this case.

Case 2

Suppose there is some $t' \geq t$ such that

$$x_k(t'|s_1, \ldots, s_{k-1}) = x_{k+1}(\phi_{k+1}(s_1, \ldots, s_{k-1}, t')|s_1, \ldots, s_{k-1}, t'). \tag{5}$$

Observe that the right hand side of equation (5) is $g_{k+1}(s_1, \ldots, s_{k-1}; t')$. Since the left hand side of equation (5) is strictly increasing in t' by part (ii) of proposition 1, and the right hand side is weakly decreasing in t' by the induction assumption, t' is the unique value that satisfies (5). By monotonicity, part (ii) of definition 1 precludes $\phi_j < t'$, and therefore, part (iii) of that definition requires $\phi_j = t'$. This shows uniqueness. For existence, we again show that ϕ_j satisfies parts (i) - (iv) of definition 1. (i) does not apply, and (ii) and (iii) are satisfied due to $\phi_j = t'$. Finally, (iv) does not apply due to part (v) of proposition 1.

(ii) As just shown, if there exists $t' \geq t$ such that

$$x_k(t'|s_1, \ldots, s_{k-1}) = g_{k+1}(s_1, \ldots, s_{k-1}; t'), \tag{6}$$

then $\phi_k(s_1, \ldots, s_{k-1}; t) = t'$, and $\phi_k(s_1, \ldots, s_{k-1}; t) = t$ otherwise.

Observe that by the induction assumption, $t \geq s_{k-1}$. Therefore, $g_k(s_1, \ldots, s_{k-2}; s_{k-1}) = x_k(\phi_k(s_1, \ldots, s_{k-1}; s_{k-1})|s_1, \ldots, s_{k-1})$. Hence, a direct application of lemma 7 yields that $g_k(s_1, \ldots, s_{k-2}; s_{k-1})$ is continuous and weakly decreasing in all arguments.

□

D.4 Proof of proposition 2

(i) We show that σ is an SPE by arguing that no player has an incentive to unilaterally deviate to any other strategy.

Fix an arbitrary history H_t in which k employers have exited, and let $s_1, \ldots, s_{k-1}$ be the times of exit. Fix an arbitrary employer j who has not yet exited in history H_t. Define for all $t' > t$

$$y_j(t'; H_t) \quad := \quad x\Big(t' \Big| s_1, \ldots, s_{k-1}, \big(t_j(H_t)\big)_{t' > t_j(H_t), j \geq k}\Big)$$

Note that $\big(t_j(H_t)\big)_{t' > t_j(H_t), j \geq k}$ is the vector of all hiring times for history H_t that lie in the interval $[s_{k-1}, t')$.

Intuitively, $y_j(t'; H_t)$ is the payoff (student skill) that employer j obtains from taking an action that causes him to exit at time t', given the remaining employers' strategies.[61]

We prove the proposition in two steps.

First we show that other employers react to a deviation by employer j only in immaterial ways. Regardless of the strategy that employer j adopts, if, conditional on the other players' strategies σ_{-j}, this strategy makes employer j exit at time t', then j's expected payoff from this strategy is at most $y(t', H_t)$. (This is not a priori obvious, since a failure by employer j to exit at the point in time at which he would exit under σ_j could, in principle, lead to subgames in which employers other employers behave in a way that increases employer j's expected payoff.) Second, we show that the shape of

[61] Observe that $y_j(t'; H_t)$ does not take into account that an employer more desirable than j may play e_1 at some candidate hiring time $t_j(H_t)$. In this case, employer j does not obtain $y_j(t_j(H_t); H_t)$ from playing an action that causes him to exit at $t_j(H_t)$, but instead obtains $\lim_{t' \searrow t_j(H_t)} y_j(t'; H_t)$, which is strictly smaller than $y_j(t_j(H_t); H_t)$ by part (v) of proposition 1. Hence, $y_j(t'; H_t)$ is an upper bound on the payoff that player j can obtain by exiting at time t'.

y is such that employer j does not have any incentive to deviate to any other strategy, including mixed strategies.

Lemma 8. *Fix a history H_t. Suppose employer j deviates to a continuation strategy σ'_j. Let $F_j : [t, T] \to [0, 1]$ be the distribution of the time at which employer j exits in strategy profile (σ'_j, σ_{-j}) in the continuation game. Then, employer j's expected payoff satisfies*

$$\pi_j(\sigma'_j, \sigma_{-j}) \leq \int_t^T y(t'; H_t) dF_j(t').$$

Proof. The strategies in σ satisfy the following two properties by construction. First, for any history H_t, at each point in time t' at which there are $l \geq 1$ hiring times, at least $l + 1$ employers take some action $e \in E$ such that l employers exit at time t', and no employer takes any action $e \in E$ at times t' that do not coincide with any candidate hiring time. Hence, exit times depend on employer j's strategy only through the candidate hiring times it induces. Second, candidate hiring times depend solely on the points in time at which previous employers have exited. Conditional on exit times, they do not depend on what action was taken by which employer. Hence, exit times depend on employer j's strategy only if j exits at a point in time that is not a candidate hiring time for the given history. Consequently, employer j's payoff from exiting at t' with probability 1 is given by $y(t'; H_t)$ if no more desirable employer exits simultaneously. If a more desirable employer exits simultaneously, then by proposition 1, part (v), employer j's payoff is strictly less than $y(t'; H_t)$. Hence, the claim follows by taking the expectation of $y(t'; H_t)$ with respect to F_j. □

Lemma 9.

(a) $y(t'; H_t)$ has a relative maximum at each $t_j(H_t)$ with $t_j(H_t) > t$ and $j \geq k$

(b) For all $j \geq k$ with $t_j(H_t) = t$, we have $x\big(t_j(H_t)|s_1, \ldots, s_{k-1}, t_k(H_t), \ldots, t_{j-1}(H_t)\big) \geq \sup_{t'>t} y(t'; H_t)$

(c) $y(t'; H_t) = y(t''; H_t)$ for all $t', t'' \in \{t_j(H_t) > t : j \geq k\}$

Proof. (a) is an immediate implication of parts (ii) and (v) of proposition 1. (b) is due to part (iv) of the definition of candidate hiring times, definition 1, and (c) follows immediately by part (iii) of that definition. □

By lemma 8, analyzing the properties of y is sufficient for bounding the payoff from a given deviation strategy σ'_j from above. By parts (a) and (b) of lemma 9, any strategy σ'_j whose support is not a subset of $\{t_k(H_t), \ldots, t_J(H_t)\}$ is strictly dominated by σ_j. If in strategy profile σ and after history H_t employer j exits at time t, then by part (b) of lemma 9, employer j has no incentive to deviate to any other strategy.[62] If under strategy profile σ and history H_t employer j exits at time $t_{j'}(H_t) > t$, then employer j obtains expected payoff $y(t_{j'}; H_t)$. Hence, by part (iii) of lemma 9, employer j has no incentive to deviate to any other strategy. This completes the proof.

(ii) Fix a (possibly mixed) SPE profile σ. We prove by induction that for any history H_t in which $k-1$ employers have exited, the kth employer exits at the candidate hiring time $t_k(H_t)$ for this history with probability 1. We anchor the induction at $k = J$, and step from k to $k-1$. (Hence, the induction assumption implies that for any history H_t in which $k-1$ employers have exited, all subsequent employers exit at the candidate hiring times $t_{k+1}(h_t), \ldots, t_J(h_t)$ for this history.)

To anchor the induction, suppose $k = J - 1$. Then, because $x_J(t|h_t)$ is strictly increasing in t, in any SPE profile the Jth student must be hired at T, which is the Jth candidate hiring time for any history.

The meat of the argument lies in the induction step, which proceeds roughly as follows. Let H_t be a history in which $k-1$ employers have already exited up until, and including time t. We first show that no employer will exit strictly before the next candidate hiring time (since at any such point in time, remaining in the game is preferable to exiting). We then suppose that with some probability, the kth exit occurs strictly after $t_k(H_t)$. Since exiting is preferable to remaining in the game at all such points in time, employers must put full probability on exiting in the interval of time during which the probability that the kth employer has not yet exited is positive. This implies that the support of each employers' exit-distribution must have the same supremum $\bar{t} > t_k(H_t)$. We then show that some employer has an incentive to deviate to a strategy in which he will have exited with probability 1 at some point strictly before $\bar{t}$. (If the strategies are deterministic, this is the incentive to undercut the competing employers' hiring times.) This contradicts the fact that each employer's exit-distribution must have the

[62] By definition, the current time t is a candidate hiring time only if exiting immediately is weakly better than remaining in the game.

same supremum. Therefore, we conclude that $\bar{t} = t_k(H_t)$, and thus, that at least one employer exits with probability 1 at candidate hiring time $t_k(H_t)$

Formally, fix a history H_t with $k-1$ previous hiring times $h_t = (t_1, \ldots, t_{k-1})$ with $k \leq J-1$.

First, we show that if $t > t_{k-1}$, then no employer may put positive probability on exiting in the time interval $(t, t_k(H_t))$. For any history H_t' with previous hiring times h_t', let $g_k(t'|h_t')$ denote an employer's expected payoff from remaining in the continuation game if all students are hired at candidate hiring times in the continuation game. Formally, this is defined by equation (3). As shown in lemma 6, $x(t'|h_t) - g_k(t';h_t)$ is strictly monotonically increasing, and by definition of $t_k(H_t)$, $x(t_k(H_t)|h_t) = g_k(t_k(H_t);h_t)$. Hence, for all $t' < t_k(H_t)$, we have $x(t'|h_t) < g_k(t';h_t)$, so that exiting at any such t' cannot be a best response for any employer.

Second, we show that at least one employer j must put probability 1 on exiting at $t_k(H_t)$.

For any employer j we let $F_j(t')$ denote the probability that employer j exits in the interval $[t, t']$. (When we subsequently alter some employer j's strategy, we will do so by altering F_j. Which actions the employer will take to achieve this is defined implicitly, from F_j and the remaining employers' strategies.) Moreover, let $\tilde{J}$ denote the set of employers who are still in the game at t.

Consider an arbitrary employer $j \in \tilde{J}$. Let $\tilde{Q}_j(t')$ denote the probability that any employer other than j exits in the time interval $[t, t']$. $\tilde{Q}_j$ is given by

$$\tilde{Q}_j(t') = 1 - \prod_{j' \in \tilde{J}, j' \neq j} (1 - F_{j'}(t'))$$

Note that $\tilde{Q}_j$ may be discontinuous.

If employer j and some other employer j' play e_1 at the same point in time t' with positive probability, the outcome for player j will depend on whether j' is more or less desirable than j. Thus, it is useful to define $\bar{Q}_j(t')$ as the probability that employer j has been preempted by a more desirable employer at or before time t'. This c.d.f may also be discontinuous.

$$\bar{Q}_j(t') = \left[1 - \prod_{j' \in \tilde{J}, j' < j} (1 - F_{j'}(t'))\right]$$

This allows us to define $Q_j(t')$ as the probability that employer j is preempted by any employer within the time interval $[t, t')$, or is preempted by a more desirable employer at time t', as follows:

$$Q_j(t') = \lim_{t'' \nearrow t'} \tilde{Q}(t'') + \left(\bar{Q}_j(t') - \lim_{t'' \nearrow t'} \bar{Q}_j(t'')\right)$$

As shown just above, no employer exits strictly before $t_k(H_t)$. So for all $t' < t_k(H_t)$ we have $\tilde{Q}_j(t') = 0$. If the probability that some employer exits at $t_k(H_t)$, which is the case if $\tilde{Q}_j(t_k(H_t)) = 1$, then the claim is proved. Hence, suppose $\tilde{Q}_j(t_k(H_t)) < 1$. We now derive employer j's expected payoff (student skills) from exiting at $t' \in [0, T]$. The probability that employer j manages to be the next employer to exit at the given time t', equals the probability that no other employer exits in $[t, t')$ or no more desirable employer exits at t'. This is given by $(1 - Q_j(t'))$. In this case, employer j obtains expected payoff $x_k(t'|h_t)$. Instead, if j is preempted by another employer, the expected payoff from his continuation strategy depends on the point in time s at which he was preempted. By the induction assumption, students are hired at candidate hiring times in the continuation game, so that employer j's expected payoff in the continuation game is given by $g_k(s; h_t)$. $Q_j(t')$ is the probability that employer j will be preempted by some other employer in the interval $[t, t']$, and $\int_0^{t'} g_k(s; h_t) \frac{1}{Q_j(t')} dQ_j(s)$ is employer j's expected payoff conditional on this event. Summing up, employer j's expected payoff from playing e_1 at t' is given by

$$\psi_j(t') = \big(1 - Q_j(t')\big) x_k(t'|h_t) + Q_j(t') \int_0^{t'} g_k(s; h_t) \frac{1}{Q_j(t')} dQ_j(s) \quad (7)$$

Hence, employer j's expected payoff from a strategy with c.d.f. of playing e_1 given by $F_1^j(t')$ (and $F_i^j(t') = 0$ for all $i > 1$) is

$$\Psi_j(t) = \int_t^T \psi(t') dF_1^j(t') \quad (8)$$

Now observe that best-responding requires the following lemma. (For any cumulative distribution function G defined on $[t, T]$, we let $\text{supp}(G)$ denote the support of G, and we let $\sup(\text{supp}(G))$ denote the supremum of this set.)

Lemma 10. *For all $j', j'' \in \tilde{J}$ we have $\sup(\text{supp}(F^{j'})) = \sup(\text{supp}(F^{j''}))$. By extension, for all $j \in \tilde{J}$,*

$$\sup(\text{supp}(F^j)) = \sup(\text{supp}(Q_j)) \tag{9}$$

Proof. Suppose w.l.o.g. that $\sup(\text{supp}(F_{j'})) > \sup(\text{supp}(F_{j''}))$ for some $j', j'' \in \tilde{J}$. This means that employer j' allocates positive probability mass to hiring the kth student to a time interval during which some other employer j'' has already hired that student with probability 1. At each time t' in this interval, $Q_j(t') = 1$, so that $\psi_j(t') = \int_0^{t'} g_k(s; h_t) dQ_j(s)$. This is strictly smaller than $x(t_k(H_t)|h_t)$, because for all $t' > t_k(H_t)$ we have $x(t'|h_t) > g_k(t'; h_t)$. Hence, employer j' could do strictly better by reallocating that probability mass to time $t_k(H_t)$. □

Let $\bar{t} = \sup(\text{supp}(Q_j))$.

Lemma 11. *F^j is continuous at $\bar{t}$ for all employers $j \in \tilde{J}$.*

Proof. Recall that for all $t' > t_j(h_t)$ we have $x_k(t'|h_t) > x_k(t_k(H_t)|h_t) = g_k(t_k(H_t); h_t) > g_k(t'; h_t)$. Because $\frac{Q_j(\cdot)}{Q_j(t')}$ is a well-defined c.d.f., we thus obtain for all $t' > t_k(H_t)$:

$$x_k(t'|h_t) > \int_0^{t'} g_k(s; h_t) d\left(\frac{Q_j(s)}{Q_j(t')}\right) \tag{10}$$

By definition of $t_k(H_t)$, and because $Q_j(t') = 0$ for all $t' < t_k(H_t)$ (as shown previously) we have $\psi_j(t_k(H_t)) = x_k(t_k(H_t)|h_t)$.

For the least desirable remaining employer j, $Q_j(\bar{t}) = \tilde{Q}(\bar{t})$. By definition of $\bar{t}$, we have $Q_j(\bar{t}) = 1$, so that $\psi(\bar{t}) = \int_0^{\bar{t}} g_k(s; h_t) d\left(\frac{Q_j(s)}{Q_j(t)}\right)$. By the assumption that $\bar{t} > t_k(H_t)$, and by the fact that $x_k(t'|h_t) > g_k(t'; h_t)$, we derive that $\psi(t_k(H_t)) > \psi(\bar{t})$. Hence, F^j, which is a best response to Q_j, must not place any positive probability on $\bar{t}$, as such a strategy could be improved be reallocating that probability to $t_k(H_t)$ instead. Hence, because $\sup(\text{supp}(F^j)) = \bar{t}$, as previously derived, and because $\bar{t} > t_k(H_t)$, F^j

must be continuous at $\bar{t}$ for the least desirable remaining employer $j \in \tilde{J}$. Therefore, for the next-to-least desirable remaining employer j', we obtain that $Q_{j'}(\bar{t}) = \tilde{H}(\bar{t})$, and can apply the same argument as above to show that $F^{j'}$ must be continuous at $\bar{t}$. Inductively, we derive that F^j must be continuous at $\bar{t}$ for all remaining employers $j \in \tilde{J}$. □

From lemma 10 we have that $\sup(\mathrm{supp}(F^j)) = \bar{t}$ for all j. Moreover, by lemma 11, $\mathrm{supp}(F^j)$ must contain $(\bar{t} - \delta, \bar{t})$ for some $\delta > 0$ (i.e. rightmost connected subset of $\mathrm{supp}(F^j)$ is a non-degenerate interval). Because F^j is a best response, its support must be a subset of maximizers t' of $\psi_j(t')$. Let $\hat{t}$ be the supremum of the set of these maximizers, i.e. $\hat{t} = \sup(\mathrm{argmax}_{t' \in [t,\bar{t}]} \psi(t'))$. Trivially, $\max_{t' \in [t,\bar{t}]} \psi_j(t') \geq \psi_j(t_j(h_t))$. Moreover, as shown before,
$\psi_j(t_j(h_t)) > \psi_j(\bar{t})$. By lemma 11, Q_j is continuous at $\bar{t}$ for all j, and therefore, ψ_j is continuous at $\bar{t}$. Because ψ_j is continuous at $\bar{t}$, and due to $\psi_j(t_j(h_t)) > \psi_j(\bar{t})$, we therefore conclude that there exists $\epsilon > 0$ such that $\psi_j(t') < \psi_j(t_j(h_t))$ for all $t' \in (\bar{t} - \min\{\epsilon, \delta\}, \bar{t})$. Consequently, $\sup(\mathrm{supp}(F^j)) \leq \hat{t} < \bar{t}$. This, however, contradicts lemma 10.

This finishes the proof of the induction step, and thus the proof of proposition 2.

We note that the above proof has the following immediate implications:

- No more than one employer can exit at a given hiring time $t_k(H_t) > t$. This is because if two employers exited, then the second would be exiting strictly before the next hiring time, in contradiction to what was just proved.
- Employers participate in hiring frenzies exactly as part (iii) of definition 1 requires: In case of a hiring frenzy, there are simply multiple hiring times at the same time t.

D.5 Proof of corollary 1.

(iii) We construct an equilibrium in which each student receives multiple simultaneous offers at each equilibrium hiring time, and in which the order of exit is $(j_1, \ldots, j_{J-1}, J)$ where $(j_1, \ldots, j_{J-1})$ denotes an arbitrary permutation of the vector $(1, \ldots, J-1)$. We alter part (iii) of the definition of equilibrium strategies as follows: $\sigma_j(H_t) = e_m$ if

$t_{k+1}(H_t) = t_{k+m}(H_t) = t$ and if $x(t_{k+m}(H_t)|t_1, \ldots, t_k, t_{k+1}(H_t), \ldots, t_{k+m-1}(H_t))$ $> x(T|t_1, \ldots, t_k, t_{k+1}(H_t), \ldots, t_{J-1}(H_t)) + \epsilon$. Each employer plays an equilibrium strategy with this alteration, and the ambiguity in part (iv) is resolved as follows: at the kth equilibrium hiring time, employers j_k and j_{k-1} play e_1. All other employers play s. At each hiring time that is off the equilibrium path, all employers play e_1. This strategy profile constitutes a subgame perfect ϵ equilibrium in which the order of exit on the equilibrium path is given by $(j_1, \ldots, j_{J-1})$. Since this argument holds for all $\epsilon > 0$, it holds in the limit. The proof for the fact that a unique order of exit is consistent with exact equilibrium is shown in the main text.

(iv) We prove this corollary by induction over j. To see the base case, suppose that $t'_1 \geq t_1$. By proposition 2, in equilibrium all employers obtain a student with the same expected skill level. In particular, $x_J(T|t'_1, \ldots, t'_{J-1}) = x_1(t'_1)$. By the assumption that $t'_1 \geq t_1$ and because $x_1(t)$ is strictly increasing, $x_1(t'_1) \geq x_1(t_1)$. Because $x_2(t|t_1)$ is strictly increasing in the first argument, and decreasing in the second, $x_2(t'_2|t'_1) = x_1(t'_1) \geq x_1(t_1) = x_2(t_2|t_1)$ implies $t'_2 \geq t_2$. Inductively we thus find that $t'_j \geq t_j$ for all $j \leq J-1$. Because $x_J(T|t_1, \ldots, t_{J-1})$ is decreasing in $t_1, \ldots, t_{J-1}$, we thus obtain that $x_1(t'_1) \geq x_J(T|t'_1, \ldots, t'_{J-1}) > x_{J+1}(T|t'_1, \ldots, t'_J)$. The last inequality is due to part (v) of proposition 1. This contradicts the fact that all hiring times yield the same expected student skill level.

For the induction step, suppose that $t'_l < t_l$ for all $1 \leq l \leq j-1$. Again, we use proof by contradiction. Assume $t'_j \geq t_j$. Then, $x_1(t'_1) = x_j(t'_j|t'_1, \ldots, t'_{j-1}) \geq x_j(t_j|t'_1, \ldots, t'_{j-1}) \geq x_j(t_j|t_1, \ldots, t_{j-1}) = x_1(t_1)$, where the equalities follow by proposition 2, the first inequality is because x_j is increasing in the first argument, and the second is due to the induction assumption and the fact that x_j is decreasing in all other arguments. This contradicts the fact that x_1 is strictly increasing.

D.6 Proof of proposition 3

To simplify notation, we summarize a vector of the form $(t_0, \ldots, t_0, T, \ldots, T)$ with k entries t_0 and k' entries T by (k, k'). We let $x_{k+1}(t_0|k) = x_{k+1}(t_0|t_0, \ldots, t_0)$, for $k, k' \in \{1, \ldots, J\}$. Hence, $x_{k+k'+1}(T|k, k') = x_{k+k'+1}(T|t_0, \ldots, t_0, T, \ldots, T)$, where the vector $t_0, \ldots, t_0, T, \ldots, T$ in the foregoing expression contains k entries t_0 and k' entries T.

(i) First, we show that any equilibrium strategy profile is in threshold strategies.

Fix an arbitrary pure strategy SPE profile σ.

We show that there is an employer $\hat{j} \in \{1, \dots, J\}$ such that the $\hat{j}$ most desirable employers $j \le \hat{j}$ exit at T, and the remaining employers exit at t_0. (Observe that this implies that expected equilibrium payoffs are weakly higher for more desirable employers, $\pi_1 > \dots > \pi_{\hat{j}} = \pi_{\hat{j}+1} = \dots = \pi_J$.)

Suppose this is not the case. We show that there exists an employer who has an incentive to deviate. Let k be the number of employers who hire at t_0. Consider the most desirable employer j_1 who hires at t_0. From hiring at t_0, this employer obtains payoff $z_{j_1}(t_0) = x_1(t_0) = 0$. If he deviated to hiring late, this employer would obtain $x_{k+j_1-1}(T|k-1, j_1-1)$. This is because if j_1 hires at T, $k-1$ students will have been hired at t_0, and $j_1 - 1$ more desirable employers will compete with j_1 at T (because j_1 is the most desirable employer who hires early). Because σ is an equilibrium profile, j_1 cannot profit from this deviation. Hence, $x_{k-1+j_1}(T|k-1, j_1-1) \le x_1(t_0) = 0$. Now consider the best employer $j_2 > j_1$ who is less desirable than j_1 and who hires at T. By doing so, he obtains at most $x_{k+j_1}(T|k-1, j_1)$. By part (v) of proposition 1, $x_{k+j_1}(T|k, j_1-1) < x_{k-1+j_1}(T|k-1, j_1-1) \le x_1(t_0) = 0$. By hiring at t_0 instead, j_2 can obtain 0.

It remains to show that a pure strategy equilibrium profile exists. For any $k \in \{0, \dots, J-1\}$, define $y(k) = x_{k+1}(T|k, 0)$. This is the expected payoff of the least desirable employer who exits at T, when k employers exit at t_0. Because $x_J(t_0|J-1) = 0$, and because $x_J(t|J-1)$ is strictly increasing in t by part (i) of proposition 1, we have $y(J-1) = x_J(T|J-1, 0) > 0$. Hence, there exists $k \in \{0, \dots, J-1\}$ such that $y(k)$ is nonnegative. Let k^* be the smallest such number. Then define the strategy profile σ as follows: Any sufficiently desirable employer $j \le J - k^*$ exits at T and all remaining employers exit at t_0. It follows by the definition of $y(k^*)$ that no employer has an incentive to deviate.

Finally, it is immediate that if $x_J(T|J-1) > 0$, then the unique equilibrium is for all employers to exit at T.

(ii) To show that there is no pure strategy equilibrium if t_0 is sufficiently close to T, we first consider the case in which student qualities do not change over time. Fix a pure strategy profile σ. Let J_1 and J_2 be the sets of employers that hire at t_0 and T, respectively. For all $j = 1, \ldots, J$, let $y_j := x_j(T|0, j-1)$ be the payoff from hiring the jth best student at T. Then, the k-most desirable employer in set J_1 obtains expected payoff y_k. Similarly, the k'-most desirable employer in set J_2 obtains $y_{|J_1|+k'}$. Thus, any employer in J_1 earns at least as much as any employer in J_2.

In this case, there are at most two pure strategy equilibria: (1) all employers hire at t_0, and (2) all but the least desirable employer J hire at t_0, and employer J hires at T. We show that no other strategy profile can be an equilibrium.

Let $\hat{j}$ denote the least desirable employer who hires early in profile σ, and suppose $\hat{j} \leq J-2$. Then, employer $\hat{j}$ earns at least y_{J-2}. Moreover, employer J and $J-1$ hire at T, and earn y_J and y_{J-1}, respectively. By deviating to hiring at t_0, employer J can improve his payoff to at least y_{J-1}.

Second, we consider the case in which $t_0 < T$ so that student skills do change over time. Note that in the first case, we have derived strictly positive deviation incentives for all but two strategy profiles σ. Hence, because there are finitely many strategy profiles, and because the conditional expectation functions are continuous by part (i) of proposition 1, standard topological arguments show that there exists $\eta > 0$ such that for all $t_0 \in (T-\eta, T)$, for any strategy profile σ for which some player had an incentive to deviate when student skills do not change over time, there is some player who has an incentive to deviate when they do change.

Hence it remains to show that it neither is a pure strategy equilibrium for all employers to hire early, nor for all but the least attractive employer to hire early. If all employers hire at t_0, then employer J obtains $x_J(t_0|J-1)$. By deviating to hiring at T, employer J obtains $x_J(T|J-1, 0)$ which is strictly profitable by part (i) of proposition 1. If all but the least attractive employers hire early, then employer $J-1$ obtains $x_{J-1}(t_0|J-2)$. By deviating to hiring at T, employer $J-1$ obtains $x_{J-1}(T|J-2, 0)$ which is strictly profitable, again by part (i) of proposition 1.

Thus, for all $t_0 \in (T-\eta, T)$ there is no pure strategy equilibrium.

Finally, we show that if t_0 is sufficiently close to T, then in any equilibrium, the employers' expected equilibrium payoffs satisfy $\pi_1 > \pi_2 > \ldots > \pi_J$. Consider any employer $k \in \{1, \ldots, J\}$. From hiring at t_0, employer k will obtain at least

$$\pi_k > x_k(t_0|k-1) \tag{11}$$

(since the previous expression is strictly decreasing in k by part (v) of proposition 1, and since $k-1$ employers are more desirable than employer k). This provides a lower bound on employer $\tilde{j}$'s expected equilibrium payoff.

Moreover, the sum of payoffs to employers cannot exceed π_{max}, defined by $\pi_{max} = \sum_{j=1}^{J} x_j(T|T, \ldots, T)$. Hence, we bound employer k's payoff from above by

$$\begin{aligned} \pi_k &< \pi_{max} - \sum_{j=1, j\neq k}^{J} x_j(t_0|j-1) \\ &= x_k(T|0, k-1) + \sum_{j=1, j\neq k}^{J} \left[x_j(T|0, j-1) - x_j(t_0|j-1)\right] \end{aligned} \tag{12}$$

As t_0 approaches T, the second term in expression (12) approaches 0, by continuity of x_j. Moreover, the left hand side of (11) approaches $x_k(T|0, k-1)$. Consequently, the payoffs in any equilibrium approach the payoffs from an assortative matching at T. Hence, there exists $\delta > 0$ such that for $t_0 \in (T - \delta, T)$ we have $\pi_1 > \pi_2 > \ldots > \pi_J$ in any equilibrium, as was to be shown.

Hence, part (ii) of proposition 3 holds for all $t_0 \in (T - \min\{\eta, \delta\}, T)$, which completes the proof.