

Equilibrium Information Disclosure and Unraveling*

Michael Ostrovsky[†] Michael Schwarz[‡]

January 4, 2003

Abstract

This paper explores information disclosure in matching markets, e.g. the informativeness of transcripts given out by universities. We show that the same amount of information is disclosed in all equilibria. We then demonstrate that if universities disclose the equilibrium amount of information, students and employers will not find it profitable to contract early; if they disclose more, unraveling will occur.

Keywords: matching markets, information disclosure, early contracting, unraveling, grade inflation.

*We are grateful to Drew Fudenberg, Kate Ho, Ed Lazear, Muriel Niederle, Al Roth, Andrei Shleifer, Larry Summers, Adam Szeidl, and seminar participants at Brandeis University, Harvard University, Virginia Polytechnic Institute, and the NBER Labor Studies Meeting for comments and suggestions.

[†]Department of Economics, and Graduate School of Business Administration, Harvard University, Littauer Center, Cambridge, MA 02138 (mostrovsky@hbs.edu)

[‡]Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138 (mschwarz@fas.harvard.edu)

When recruiters call me up and ask for the three best people, I tell them, ‘No! I will give you the names of the six best.’

Professor Robert J. Gordon, Director of Graduate Placement,
Northwestern University, Department of Economics

Harvard wants high schools to give class rank, but high schools do not want to.

Senior Harvard official

1 Introduction

In a variety of situations, sellers choose to withhold some information about their goods from buyers. This is easy to explain if buyers can be fooled by the lack of information and hold erroneous beliefs about the goods. However, the lack of rationality, or the lack of common knowledge, are not necessary conditions for incomplete information disclosure. In this paper we model information disclosure in matching markets and show that noisy signals can be a necessary feature of a fully rational equilibrium.

Let us start with examples. The first one is the phenomenon of grade inflation. The words “grade inflation” suggest that the inflation of grades is similar to the inflation of currencies. The analogy is very imprecise. A period of monetary inflation merely adds zeros to prices, without making a long-term impact on the economy. Not so with grade inflation: years of grade inflation make grades permanently less informative. If half of the class gets “A”s, a lot of information is thrown out (Figure 1, while by no

means conclusive, suggests that the informativeness of grades at Harvard, as measured by their entropy, has declined in recent years as the percentage of “A”s and “A-”s has risen). This begs a question about costs and benefits to a school from making grades less informative. Clearly, less informative grades are good for bad students and bad for good students. The tradeoff, however, is not obvious.

Thus, in order to understand grade inflation we should think systematically about how informative a school would like to make its transcripts. We consider the equilibrium amount of information that schools would like to convey in their transcripts and offer a framework for estimating long-term costs and benefits of reducing their informativeness. We show that even if there is no short-run benefit from inflating grades, schools may prefer to make transcripts less than perfectly informative. Thus, reducing the informativeness of grades (by means of grade inflation) may be an optimal response to a changing environment, even if employers can not be temporarily fooled by high grades.

Another example of incomplete information disclosure is the tendency of professors recommending their students for admission to graduate schools to avoid direct comparison and rankings. Instead, they often praise different students for different skills and characteristics, making inference problematic.

Incomplete information disclosure may also arise in dynamic settings, where the timing of the disclosure influences the amount of available information. Early Action and Early Decision admission programs at many

selective colleges allow high school students to submit their applications in the Fall of their senior year, and admission decisions are made before Fall semester grades are available (in contrast to the regular admission process, which takes Fall semester grades into account). Sometimes, competing agents try to set the decision date before others, and the process “unravels”. A dramatic example of such situation is the market for law clerks (Avery, Jolls, Posner, and Roth, 2001). The interviews for clerkship positions are currently held at the beginning of the second year of three-year law school programs, when only a third of the student’s grades are available. Clearly, a lot of information is withheld. (See Roth and Xing, 1994, for examples of markets where the timing of transactions has unraveled. Avery, Fairbanks, and Zeckhauser, 2001, give a comprehensive account and analysis of early admissions programs at U.S. colleges.)

We model information disclosure in the setting where schools evaluate students and give them transcripts. Subsequently, these transcripts are used by outsiders (e.g. employers, professional schools, clerkship positions) in their hiring decisions. We assume that the ability of each student and the distribution of students among schools are given exogenously. The ability of students is perfectly observed by schools but not by outsiders. Each school decides how much information to reveal in its transcripts in order to maximize the average desirability of placement of its alumni.¹ Outsiders use transcripts

¹Even though schools are not monolithic decision makers, and grades are set by individual professors, school administrators have a variety of ways to control the informativeness of transcripts: they can introduce forced curves, publish average class grades on students’

to infer the expected ability of students and rank them solely according to their expected ability. The desirability of each position is common knowledge, and students rank positions based on desirability (thus all students have the same preferences and so do all recruiters). We assume that the supply of placement slots of a given desirability is fixed exogenously. This is a sensible assumption if we think of schools as high schools and placement positions as college slots. There is a “rent” attached to being admitted to a good college. The amount of “rent” captured by successful applicants probably has little relation to the quality of the application pool. We can also think of “schools” as colleges and placement slots as admission to professional schools², or “schools” as law schools and placement slots as clerkships in the US courts³, or in general placement slots as any jobs that have rents associated with them.

Consider the competition for admission to medical schools. Strategically introducing noise into transcripts may enable a college to substantially increase placement into moderately desirable medical schools at the cost of slightly reducing the number of students placed at top medical schools. Essentially, a college changes the distribution of placement desirability of its transcripts, reveal students’ ranks, etc.

²It is a reasonable approximation to assume that professional schools do not adjust tuition and the number of admission slots based on the ability of available applicants.

³The pay of clerks is determined exogenously. There is a more or less clear hierarchy of clerkships in terms of desirability. We can assume that the post-clerkship wages of lawyers are the sum of ability and the enhancement of human capital obtained during the clerkship. Then, the desirability of a clerkship can be defined as the enhancement of human capital obtained during the clerkship net of costs associated with it. If the ability of clerks changes, the desirability of clerkship need not change.

graduates by introducing noise into transcripts. Notice that the aggregate distribution of positions does not change as a result of noisy transcripts, and so the total desirability of placements is unchanged. However, we will see that in a broad range of situations noise is a necessary feature of equilibrium transcripts.

The relationship between the distribution of desirability and the informativeness of grades is often not transparent. For example, one may think that if the premium for excellence goes up, i.e. the world becomes more and more of a “winner takes all” environment, grades should become more informative to clearly indicate star students. However, this is not necessarily the case. Suppose over time the difference in desirability between medical schools ranked 1 and 10 increases, but the difference in desirability between medical schools ranked 10 and 20 increases even more (one may think of a medical school’s desirability as a fixed effect that it has on its students’ subsequent wages). Then the equilibrium amount of noise in transcripts should increase, because it becomes more profitable to use pooling, thus “minting” more students of the “top 10” level instead of having more students of “top 1” and “top 20” levels. The number of students in the tenth-ranked medical school does not change, but the equilibrium expected ability of students going there does.

The change in the *expected* ability of students in the tenth-ranked school is not due to the change in the distribution of *true* ability in the student population—the latter is exogenously fixed. Rather, it is due to colleges

varying the amount of noise in their transcripts, thus changing the distribution of *expected* ability. In Section 3 we show that the amount of information disclosure can be meaningfully quantified by the aggregate distribution of expected ability. The more noise the transcripts contain, the lower the variance of the distribution of expected ability is.

Theorem 1 establishes that both the equilibrium amount of information disclosure and the utility of every employer are the same in all equilibria. The equilibrium amount of information disclosure is independent of how students are distributed among colleges. It is solely determined by the aggregate distribution of student ability and the distribution of position desirability.

Lemma 2 shows that the desirability mapping in equilibrium, i.e. the function that maps student expected abilities to the desirabilities of positions they obtain, is convex. The intuition of this result is straightforward—if the mapping were concave on an interval, a college producing students of expected abilities in that interval would find it profitable to mix.

Based on Lemma 2, we then show that there is no unraveling under the equilibrium information disclosure. Intuitively, if a student finds it profitable to withhold some information by signing a contract earlier, his school would also find it profitable to exclude this information from its transcripts. Thus, if schools behave optimally and disclose the equilibrium amount of information, a student can not benefit by moving early and getting “insurance” against the arrival of additional information. Moreover, we show that if more than the equilibrium amount of information is disclosed (for example, if students

can signal their ability using test scores not controlled by schools), some students do benefit from contracting early (and some sections of the desirability mapping are concave). These results allow us to better understand when unraveling can be explained by the demand for insurance.⁴

The rest of the paper is organized as follows. In Section 2 we introduce the formal model of grade inflation. Section 3 defines equilibrium information disclosure and shows some of its properties. Section 4 considers unraveling and shows that it can not happen under equilibrium information disclosure. Section 5 concludes.

2 The Model

Consider a population of students. The ability of each student is a real number a in the interval $A = [a_L, a_H]$. Each student attends one of I schools. We assume that the distribution of ability levels at school i , represented by a Borel measure $\lambda_i(a)$, is given exogenously and is common knowledge. Without loss of generality, we assume that schools observe the true abilities of their students.⁵

Each school decides how precise to make its transcripts. A school can make transcripts completely informative, revealing the ability level of each

⁴Li and Rosen (1998), Li and Suen (2000), Suen (2000) present models of insurance-driven unraveling.

⁵Suppose nobody observes the true ability, but each school observes a signal regarding the true ability of each of its students. Based on this signal a school can form expectation about a student's ability. All results in the paper continue to hold if instead of "true ability" we use "expected ability based on information available to schools."

student, or it can make them completely uninformative, or anything in between. For example, Harvard Business School has a policy that prohibits employers from asking students about grades. The only information the school reveals is about students' honors, with about 80% of the class not receiving any.⁶

More formally, a school chooses a transcript structure, which consists of two parts. First, it chooses a set of all possible transcripts. Second, it chooses how to map the abilities of students into these transcripts. This mapping may be stochastic, i.e. for each ability a there is a probability distribution over the set of transcripts that a student of this ability can get.

Definition 1 *School i 's transcript structure TS_i is a pair $(T_i, F_i(\cdot))$, where $T_i \subseteq \mathbb{R}$ is a Borel set of possible transcripts, and $F_i(\cdot)$ is a Borel measure on $A \times T_i$ such that the marginal distribution of F_i on A equals λ_i , i.e. for any Borel set $X \subseteq A$, $F_i(X \times T_i) = \lambda_i(X)$.*⁷

After schools announce transcript structures and give transcripts, students and positions are matched. On one side of the market there is a population of students. On the other side there is a mass of positions. The desirability of each position, $q \in [q_L, q_H]$, is common knowledge. The distri-

⁶“Distinction and High Distinction are part of the MBA degree and are therefore public information. First-Year Honors and Second-Year Honors are also considered public information[...].”—HBS Honors Policy.

⁷This definition is very similar to the definition of “information structure” in Bergemann and Pesendorfer (2001). That paper, however, considers information disclosure in a very different environment—a single-seller, single-object auction, whereas we consider a matching market.

bution of desirabilities is given by a Borel measure μ . The mass of positions is equal to the total mass of students, i.e. $\mu([q_L, q_H]) = \sum_i \lambda_i(A)$.⁸

We assume that schools control the information available to employers. Based on transcripts, employers can compute the expected ability of each student.⁹ Notice that generally the distribution of *expected* ability (conditional on transcripts) will be different from the distribution of *true* ability (unless transcripts are completely informative, i.e. any two students of different true abilities get different transcripts). In the Harvard Business School example above, 80% of the resulting distribution of *expected* ability is concentrated at one point (assuming that no information aside from the transcript is available to recruiters), even though *true* ability could have any distribution.

Students rank positions by desirability, and employers rank students by expected ability. The resulting rankings induce a unique (up to permutations of equally desirable positions) assortative stable matching between students and positions. Note that the assumption that employers are risk-neutral and rank students solely by expected ability allows for a wide variety of employer payoffs, including the ones commonly assumed in the matching literature, e.g. $F = \theta(a)\phi(q)$ (Suen 2000); $F = aq$, $F = aq^2$ (Li and Suen 1998); $F = aq$

⁸This is not a restrictive assumption, because unemployment can be viewed as a position of the lowest desirability, and because if the mass of positions is greater than the total mass of students, the same subset of positions gets assigned a student under any information disclosure.

⁹We assume that schools can commit to their transcript structures. However, even if they could not, the equilibrium information disclosure would still remain an equilibrium outcome of the resulting cheap-talk game (see Crawford and Sobel (1982) for a formal analysis of cheap-talk games). On the other hand, there are many other equilibria in the resulting cheap-talk game, in some of which even more information may be disclosed.

(Haruvy, Roth, and Unver 2001).¹⁰

Each school selects a transcript structure to maximize the total desirability of positions obtained by its students. Each school is small relative to the labor market and is a “price taker”—its actions have no effect on the placement of students of a given expected ability.¹¹

The following example is an illustration of the model.

Example 1. Consider a simple setup: student abilities at each school are distributed uniformly on $[0, 100]$, and position desirabilities are distributed uniformly on $[0, 200]$. If all schools fully reveal student abilities (e.g. set $t \equiv a$), the resulting mapping from abilities to position desirabilities is linear ($Q(a) = 2a$) and no school can benefit by deviating. Thus, fully informative transcripts form an equilibrium, and, as we later show in Corollary 2, no other equilibria exist.

Let us now see what can happen out of equilibrium. Suppose for an exogenous reason all schools but one make transcripts noisy. Namely, suppose a student of ability a receives transcript t , a random draw from $U[a, a + 100]$. Then transcripts range from 0 (the least able student with the least

¹⁰This is true because as long as the output of a worker is a function of his ability, we can find a rescaling of ability such that a particular firm is indifferent between having a worker of ability a_0 for sure and a worker of uncertain ability with expectation a_0 . However, we do have to assume that this rescaling is the same for all firms. This does not preclude a possibility of complementarity between worker ability and position desirability: for example, it is consistent with our model if the output of a worker in a firm of desirability q is a product of his ability and some function of desirability, $g(q)a$, because $E_a[g(q)a] = g(q)E_a[a]$.

¹¹This can be reconciled with a finite number of schools using the standard general equilibrium approach—assume that there are I school types and an infinite number of schools of each type.

lucky draw) to 200 (the most able student with the most lucky draw). The distribution of transcripts is triangular with the peak at 100. The distribution of expected abilities, F , is also triangular, with the peak value of $\frac{1}{50}$ at 50 (figure 2).¹²

A student of expected ability \hat{a} gets a position of desirability Q , with the same rank as his own. Desirability mapping $Q(\hat{a})$ is S-shaped (figure 3).¹³

The best response of a representative school to this desirability mapping is to reveal ability truthfully if it is less than a^* and pool together all students with ability above a^* , where $a^* \approx 23$.¹⁴ Notice that the best response of a better than average school that has a distribution of student types given by $U[50, 100]$ is to make grades completely uninformative.

¹²If $t \leq 100$, the true ability could be anywhere from 0 to t (distributed uniformly on this interval, since both noise and the underlying true ability are uniform and independent) and the expected ability is $t/2$. Similarly, if $t \geq 100$, the distribution of true abilities of students with transcript t is uniform on $[t-100, 100]$, and the expected ability of a student conditional on his transcript is also given by $\hat{a}(t) = E[a|t] = \frac{(t-100)+100}{2} = \frac{t}{2}$.

¹³This student is better than $\int_0^{\hat{a}} dF$ fraction of the student population, and thus gets a position of desirability $Q(\hat{a}) = 200 \int_0^{\hat{a}} dF$. Computing the integral, $Q(\hat{a}) = \frac{\hat{a}^2}{25}$ for $\hat{a} \leq 50$ and $200 - \frac{(100-\hat{a})^2}{25}$ for $\hat{a} \geq 50$.

¹⁴To solve for the best response of U-State, notice that it will pool together students from the concave part of $Q(\hat{a})$ until the amount gained by the marginal added student a^* , $Q(\frac{a^*+100}{2}) - Q(a^*)$, equals the total amount lost by the pooled students of higher ability due to the decrease in the average pool ability, $(100 - a^*) \cdot \frac{Q'(\frac{a^*+100}{2})}{2}$. Thus,

$$\frac{Q(\frac{a^*+100}{2}) - Q(a^*)}{\frac{100-a^*}{2}} = Q'(\frac{a^*+100}{2}),$$

which implies $a^* = \frac{100}{7}(3 - \sqrt{2})$.

3 Equilibrium Information Disclosure

In our setup, the behavior of students and positions is straightforward—they get matched to the agents of the highest quality available to them on the other side of the market (in the next section we give them some flexibility by allowing early contracting). Thus, we focus on the actions of schools and the resulting transcript structures. Transcript structures form an equilibrium if no school can increase the average desirability of its students’ placement by changing its transcript structure.

The main result of this section is that in all equilibria employers of a given desirability get students of the same expected ability (even though the distributions of ability among their hires may be different). This result continues to hold regardless of the distributions of student ability within schools, as long as the distribution of ability in the entire student population remains the same.

3.1 Uniqueness of the Equilibrium Amount of Information Disclosure

In this section we show that the same amount of information is disclosed in all equilibria. We start with the definition that makes the words “amount of information disclosure” precise. Let $\omega = (TS_1, TS_2, \dots, TS_I)$ be a profile of transcript structures and let $G = \sum_i G_i$ be the distribution of expected abilities generated by ω , where G_i is the distribution of expected abilities

(conditional on transcripts) at school i . Note that if each school introduces more noise in its transcripts, the variance of the resulting distribution of expected abilities decreases.¹⁵ This leads to a natural partial ordering on the set of profiles of transcript structures.

Definition 2 *Profile of transcript structures ω is more informative than profile of transcript structures ω' if distribution G of expected abilities generated by ω is second-order stochastically dominated by distribution G' of expected abilities generated by ω' .*¹⁶

This partial ordering has two extreme elements: the completely uninformative profile, which has zero variance, and the profile revealing all student abilities, which has the highest possible variance. Also, it is clear that a more informative profile has a higher variance than a less informative one, since the former is a mean-preserving spread of the latter.

We also need to introduce the *desirability mapping*—function $Q(\hat{a})$ denoting the desirability of the position that a student with expected ability \hat{a} gets in a given equilibrium: $\forall \hat{a} \in [a_L, a_H], \mu([q_L, Q(\hat{a})]) = G([a_L, \hat{a}])$. In Section 3.2 we show that this function is well-defined. Notice that if two profiles produce different distributions of expected student abilities, the resulting

¹⁵To add noise, school i creates a new transcript structure $(T'_i, F'_i(\cdot))$ where $F'_i(t'_i|a) \equiv \int_{T_i} H'_i(t'_i|t_i) dF_i(t_i|a)$; $(T_i, F_i(\cdot))$ is the old transcript structure and $H(\cdot)$ is a distribution on $T_i \times T'_i$. If ω' is the new profile of transcript structures, and G' is the distribution of expected abilities corresponding to ω' , then G is a mean-preserving spread of G' . This is true because ω' is obtained by mixing students of different expected abilities under ω .

¹⁶Slightly abusing terminology, we will identify different transcript structures if they generate the same distributions of expected student abilities. For example, a transcript structure that gives the same grade to every student is, for our purposes, the same as a transcript structure that gives completely random grades to all students.

desirability mappings $Q_1(\hat{a})$ and $Q_2(\hat{a})$ will also be different. Thus, identical desirability mappings can only be generated by identical distributions of expected abilities in the student population.

We are now ready to state the main result of this section. In all connected equilibria, desirability mappings are the same. Equivalently, in all connected equilibria the same amount of information is disclosed. (We define “connectedness” in Section 3.2. The class of connected equilibria includes, for example, all equilibria in which at least one school produces students of all expected abilities.) The equilibrium amount of information is independent of how students are assigned to schools—only the aggregate distribution of student abilities and the distribution of position desirabilities matter.

Theorem 1 *Suppose there is a connected equilibrium with desirability mapping $Q_1(\hat{a})$. Suppose students are reshuffled among schools so that the aggregate distribution of student abilities remains the same, but the distributions of abilities within schools change, and suppose there is a new connected equilibrium with desirability mapping $Q_2(\hat{a})$. Then $Q_1(\hat{a}) \equiv Q_2(\hat{a})$.*

Proof. See Section 3.2 for auxiliary lemmas and Appendix C for the proof.

■

A straightforward corollary is that when there are two equilibria in a market, their desirability mappings coincide.

Corollary 1 *Suppose there is a market that has two connected equilibria, with desirability mappings $Q_1(\hat{a})$ and $Q_2(\hat{a})$. Then $Q_1(\hat{a}) \equiv Q_2(\hat{a})$.*

Another corollary is that if truthful revelation of abilities is an equilibrium, then there are no other connected equilibria.

Corollary 2 *Suppose there are two connected equilibria in a market, and one of them is fully informative. Then the other one also has to be fully informative, and therefore the same.*

Proof. By Corollary 1, the desirability mappings of these two equilibria have to be the same. Therefore, the distributions of expected abilities generated in these equilibria also have to be identical (since they are uniquely determined by the mapping and the distribution of position desirability). But the fully informative equilibrium is strictly more informative than any other, and so the second equilibrium also has to be fully informative, hence the same. ■

3.2 Properties of Equilibrium Information Disclosure

In this subsection we establish several lemmas describing some interesting properties of equilibrium information disclosure. These lemmas are also used to prove the main results of the paper. We assume that distributions of position desirability and student ability in each school are atomless and have full supports on closed intervals. We also assume that if all abilities are revealed truthfully, function $Q_T(a)$ mapping ability level a to position desirability Q_T of the same rank does not switch from convexity to concavity infinitely often (i.e. there exists a finite increasing sequence of ability levels a_i , starting at

the lowest and ending at the highest true ability, such that Q_T is convex or concave on each interval $[a_i, a_{i+1}]$.

We first show that in equilibrium there is a one-to-one mapping from expected ability to position desirability, i.e. students of the same expected ability get equally desirable positions. This allows us to talk about this desirability mapping as an invertible function $Q(\hat{a})$.

Lemma 1 *In equilibrium, any two students of the same expected ability \hat{a} obtain equally desirable positions.*

Proof. See Appendix. ■

Desirability mapping $Q(\hat{a})$ is monotonically increasing. This, however, does not necessarily mean that a student of a higher true ability will get matched to a better position than a student of a lower true ability: if a school gives out transcripts that are not fully informative, the lower ability student may receive a better transcript than the higher ability student and thus get a better position.

Definition 3 *An equilibrium is fair at a particular value of position desirability q if there is an ability level that is necessary and sufficient for receiving a position of this quality. More precisely, equilibrium is fair at desirability q if for any $\delta > 0$ there exist ability levels $a_1 < a_2$ and a positive $\epsilon < \delta$ such that all¹⁷ students who are hired into positions of desirabilities $[q - \epsilon, q + \epsilon]$ have ability in the interval $[a_1, a_2]$ and among the students with ability in the*

¹⁷Up to a set of measure zero.

interval $[a_1, a_2]$ nobody receives positions of desirability outside the interval $[q - \epsilon, q + \epsilon]$.

It is straightforward to show that an equilibrium is fully informative if and only if it is fair at every position desirability.

Another property that we introduce is “connectedness.” We will say that an equilibrium is connected if, roughly, there exists a sequence of schools, from best to worst, such that the first one produces some students of the highest expected ability in the population, the last one produces students of the lowest expected ability, and each school in the sequence produces a positive mass of students who are worse in expectation than the next school’s best student. The following definition formalizes this idea.

Definition 4 *Let $\hat{a}_L = \sup\{\hat{a} | Q(\hat{a}) = Q(a_L)\}$ be the lowest and $\hat{a}_H = \inf\{\hat{a} | Q(\hat{a}) = Q(a_H)\}$ the highest expected ability levels produced in an equilibrium. Then we say that the equilibrium is connected if for any expected ability $\hat{a} \in (\hat{a}_L, \hat{a}_H)$ there exist $i \in [1 \dots I]$ and $\epsilon > 0$ such that for any nonempty interval $(a_1, a_2) \subset (\hat{a} - \epsilon, \hat{a} + \epsilon)$ school i produces a positive mass of students with expected abilities in (a_1, a_2) , i.e. $G_i((a_1, a_2)) > 0$.*

Connectedness is a mild restriction. Indeed, if at least one school gives out some transcripts implying the worst and the best possible expected abilities, and everything in between, this restriction is satisfied.¹⁸

¹⁸In Appendix E we also show that a sufficient condition for the existence of a connected equilibrium is for schools to be symmetric.

Desirability mapping $Q(\hat{a})$ is monotonically increasing on $[\hat{a}_L, \hat{a}_H]$. We are now ready to characterize it further. Suppose a school produces positive masses of students with expected abilities in arbitrarily small neighborhoods of b and c . This could only be optimal for the school if by mixing students of these abilities it could not raise its payoff, i.e. if $\alpha Q(b) + (1 - \alpha)Q(c) \geq Q(\alpha b + (1 - \alpha)c)$ for any $\alpha \in [0, 1]$. Since this reasoning can be applied to every pair of points, and in a connected equilibrium there is a school producing students in a neighborhood of any point, $Q(\hat{a})$ has to be convex on $[\hat{a}_L, \hat{a}_H]$.

On the other hand, if a school does mix in any neighborhood of some desirability ν_0 , i.e. the equilibrium is not fair at that point, then it has to be the case that $Q(\hat{a})$ is concave there—otherwise the school could do better by producing students of expected abilities $\nu_0 - \epsilon$ and $\nu_0 + \epsilon$. If the equilibrium is unfair on an interval, $Q(\hat{a})$ thus has to be both convex and concave there, and therefore linear.

Conversely, if $Q(\hat{a})$ is strictly convex at a certain expected ability level \hat{a} , it is fair at $Q(\hat{a})$. Moreover, it is unprofitable for a school to mix students of abilities higher than \hat{a} with students of abilities lower than \hat{a} . Therefore, all more able students get positions better than $Q(\hat{a})$ and all less able students get positions worse than $Q(\hat{a})$. The following three lemmas summarize these results.

Lemma 2 *In any connected equilibrium, desirability mapping $Q(\hat{a})$ is convex on $[\hat{a}_L, \hat{a}_H]$.*

Lemma 3 *Suppose on some range of desirability $[\nu_1, \nu_2]$, a connected equilibrium is not fair. Then $Q(\hat{a})$ is linear in \hat{a} on $[Q^{-1}(\nu_1), Q^{-1}(\nu_2)]$.*

Lemma 4 *If $Q(\hat{a})$ is strictly convex at \hat{a} , then schools do not mix students of ability higher than \hat{a} with students of ability lower than \hat{a} .*

The next lemma shows that the lowest expected ability produced by schools is equal to the lowest true ability. This is similar to the “lowest type not signalling” in a separating equilibrium of a signalling game.

Lemma 5 *In a connected equilibrium, let \hat{a}_L be the lowest expected ability level, and a_L be the lowest true ability level. Then $\hat{a}_L = a_L$.*

Proof. See Appendix. ■

4 Unraveling

Sections 2 and 3 analyze information disclosure in a static framework. The results, however, also help better understand an inherently dynamic phenomenon that frequently occurs in matching markets—unraveling, i.e. contracting between students and positions before full information about the former is available. Examples of early contracting include Early Action and Early Decision admission programs at many selective colleges, which allow students to apply before Fall semester grades of their Senior year at high school are available (Avery, Fairbanks, and Zeckhauser 2001), the market for

federal judicial law clerks, where judges interview candidates two years prior to the beginning of the clerkships (Avery, Jolls, Posner, and Roth 2001), and many others (Roth and Xing 1994 give numerous examples of “unraveled” markets).

The literature on unraveling broadly divides the causes of early contracting into two categories: strategic considerations, or “jumping the gun” (see e.g. Roth and Xing 1994, Avery, Fairbanks, and Zeckhauser 2001), and the demand for insurance (see e.g. Li and Rosen 1998, Li and Suen 2000, Suen 2000). We focus on the insurance aspect of unraveling.

The connection between information disclosure and unraveling is close, albeit not obvious. To compute the equilibrium amount of information disclosure (as defined in this paper), one only needs to know the distribution of ability in the population and the distribution of job desirability. It is easy to check that in situations where unraveling occurs due to insurance reasons, waiting till all information is revealed will lead to the disclosure of more than the equilibrium amount of information. This is not a coincidence. Theorems 2 and 3 show that if the equilibrium amount of information is revealed, no unraveling occurs. Consequently, if schools can control the amount of information disclosed to potential employers, the insurance reason for unraveling disappears. The intuition is simple: in equilibrium, due to the convexity of desirability mapping $Q(\hat{a})$, the expected position desirability that a student will get tomorrow is higher than the position desirability that he could get today.

It may seem surprising that there is no unraveling under the equilibrium amount of information disclosure—after all, imagine all positions have similar desirability except for a few which are terrible, e.g. unemployment. Then one might think that students would be eager to sign contracts earlier to avoid this outcome. However, as the following example shows, this does not happen. What happens instead is that the equilibrium distribution of transcripts “mimics” the distribution of desirability—a small group of students gets very bad transcripts, and the rest get compressed transcripts with little information beyond being much better than the bad transcript.

Example 2. Suppose there are positions of five types. The worst position is unemployment, it has desirability 0. The remaining four positions have similar desirabilities—110, 120, 130, and 140. The share of positions of each type is 20%. At each school, ability is distributed uniformly on $[0, 100]$. First, note that it is not an equilibrium for all schools to lump all students into one category. If they do, then a school can profitably deviate by separating a small fraction of the worst students into a new category. Second, providing fully informative signals is not an equilibrium either—in this case a school can achieve the average desirability of placement of 120 by pooling all students together instead of 100 by fully separating them. Thus, schools have to partially mix in equilibrium.

The following transcript structure is an equilibrium of the market. There are five possible transcripts—we’ll call them A, B, C, D, and E, from best to worst. Entries in the following table are the probabilities of receiving a

particular transcript for students of different ability levels.

Table 1: The Probability of Receiving Transcript t Conditional on Ability Level a

Ability \ Transcript	E	D	C	B	A
$a \leq 20$	1	0	0	0	0
$20 < a \leq 56$	0	$\frac{7}{11} \cdot \frac{5}{9}$	$\frac{18}{33} \cdot \frac{4}{9}$	$\frac{15}{33} \cdot \frac{4}{9}$	$\frac{4}{11} \cdot \frac{5}{9}$
$56 < a \leq 60$	0	0	1	0	0
$60 < a \leq 64$	0	0	0	1	0
$a > 64$	0	$\frac{4}{11} \cdot \frac{5}{9}$	$\frac{15}{33} \cdot \frac{4}{9}$	$\frac{18}{33} \cdot \frac{4}{9}$	$\frac{7}{11} \cdot \frac{5}{9}$

The expected ability of a student with the worst transcript, E , is equal to 10. $\hat{a}(D) = (7 \cdot 38 + 4 \cdot 82)/11 = 54$, $\hat{a}(C) = 58$, $\hat{a}(B) = 62$, and $\hat{a}(A) = 66$. The resulting mapping $Q(\hat{a})$ is linear: $Q(\hat{a}) = \frac{5}{2}(\hat{a} - 10)$, and so a school can not increase the average desirability of placement by deviating from this transcript structure.

Notice that in this equilibrium there is no unraveling (or, more precisely, no incentive to unravel) since students become effectively risk-neutral. Consider a student whose first-year transcript indicates an expected ability level corresponding to a particular job desirability. This student can secure a job corresponding to his current expected ability or he can wait for second-year grades. In the absence of private information about ability the expected change in ability implied by the transcript must be zero. It is easy to see that the expected change in position desirability can not be negative as a result of arrival of new information.

Even if students did have private information, unraveling would still not

occur. In fact, the result would become even stronger. In the absence of private information, unraveling is a matter of indifference for both students and positions. If students do have private information, adverse selection works against unraveling, because the lowest ability students have higher payoff from unraveling than observationally equivalent students of higher ability. Essentially, only the lowest ability students are eager to unravel, and unraveling can not occur under equilibrium information disclosure except for a set of measure zero.¹⁹

In the remainder of the section, we first present a simple two-period model where no information is available in period 1, which is very similar to the model of Suen (2000).²⁰ This similarity brings into focus the fact that the schools' ability to control information undermines the insurance reason for unraveling. We then move to a more general framework, where new information arrives continuously. We show a striking result—contracting at the last possible moment (“no unraveling”) is an equilibrium only if the transcript structures eventually revealed by the schools form an equilibrium. It is not an equilibrium if more information is revealed.

¹⁹It is straightforward to make the above statement entirely formal.

²⁰We should note that Suen's model is more complicated—it involves wages. However, the main intuition that unraveling is caused by workers' demand for insurance can be applied to our model just as well, as we will show at the end of this section when we demonstrate unraveling in environments where schools cannot fully control information.

4.1 Two-period Model, No Unraveling under Equilibrium Information Disclosure

Suppose students stay in school for 2 periods. In period 1 no information about them is known. Therefore, for all students in school i expected ability in period 1 is the same, \hat{a}_i . A student has no private information about his ability²¹. Suppose employers and students can sign binding contracts in either year of study based on the information available at that period.

Theorem 2 *If schools' transcript structures in period 2 form a connected equilibrium, then no position can increase the expected ability of its match by making an early offer.*

Proof. Take a student from school i in period 1. His expected ability in period 1 is \hat{a}_i . If he waits until period 2, he will get a position of desirability $Q(\hat{a}_2)$. By the law of iterated expectations, $E_i[\hat{a}_2] = \hat{a}_i$. Desirability mapping $Q(\hat{a}_2)$ is convex, and therefore $E_i[Q(\hat{a}_2)] \geq Q(\hat{a}_i)$. Thus, a student will only accept an early offer from a position that is at least as desirable as $Q(\hat{a}_i)$. But positions of desirability $Q(\hat{a}_i)$ and higher get a student of expected ability at least \hat{a}_i if they wait until period 2, and so can not benefit from moving early.

■

²¹Allowing for private information does not change the result, except that the 0-mass of students and firms who are indifferent between contracting in periods 1 and 2 is of lower quality due to adverse selection.

4.2 Continuous-Time Model of Information Arrival

We now set up a continuous-time model of information arrival, and show a close connection between unraveling and equilibrium information disclosure.

Students are in school from time $\tau = 0$ until time $\tau = \bar{\tau}$. At time 0 no information about a student is known except for the school that he attends. While the student stays in school, new information arrives continuously and is added to his transcript (we assume that information about students can not disappear). Namely, at each time τ a potential employer can compute the student's expected ability \hat{a}_τ based on the current transcript. Since employers use Bayes' rule to form beliefs about a student's expected ability, the drift term must be zero and the process is a martingale. Consequently, \hat{a}_τ follows a diffusion process

$$d\hat{a}_\tau = \sigma(\cdot)dz_\tau, \tag{1}$$

where we assume $\sigma(\cdot)$ to be a bounded continuous function of \hat{a}_τ and τ , such that the process does not leave the interval $[a_L, a_H]$ and for all $\hat{a} \in (a_L, a_H)$ $\sigma(\hat{a}, \bar{\tau}) > 0$. We also assume that for some $\tau < \bar{\tau}$ function $\hat{Q}(\hat{a}_{\tau'}, \tau') = E[Q(\hat{a}_{\bar{\tau}})|\hat{a}_{\tau'}]$ is twice continuously differentiable for all $\tau' \in [\tau, \bar{\tau}]$.²² Whenever expected ability follows such diffusion process we will say that *information arrives gradually*.

Each position's desirability is common knowledge throughout the game, and any student-position pair can enter into a binding match any time during

²²More formally, $\hat{Q}(\hat{a}_{\tau'}, \tau')$, is twice continuously differentiable on the set of points $\{(a, \tau')|\tau' \in [\tau, \bar{\tau}], a \text{ is in the domain of } \hat{Q}(\cdot, \tau')\}$.

the game. Unraveling occurs if at some time $\tau < \bar{\tau}$ there is a pair student (S) and position (P) that finds it profitable to sign such a contract.²³

We now claim that it is an equilibrium for students and firms to sign contracts at time $\bar{\tau}$ without unraveling if transcript structures generated by schools on day $\bar{\tau}$ form an equilibrium. If more than the equilibrium amount of information is disclosed, some students and employers will find it profitable to sign contracts earlier. The intuition for the first statement is the same as in Theorem 2. The intuition for the second statement is that if more information is revealed, some portion of the resulting mapping $Q(\hat{a}_{\bar{\tau}})$ will be concave, thus making some students effectively risk-averse.

Theorem 3 *Suppose that information about ability of students arrives gradually (see equation (1)). If at time $\bar{\tau}$ transcripts contain the equilibrium amount of information, then it is an equilibrium for all students and positions to wait until time $\bar{\tau}$ to sign contracts. If at time $\bar{\tau}$ transcripts contain more than the equilibrium amount of information, then some agents are strictly better off not waiting till time $\bar{\tau}$ to sign contracts.*

Proof. See Appendix. ■

Our arguments rely on the assumption that not only ordinal, but also cardinal preferences of schools and students over positions are the same. If

²³It is profitable for the pair to sign such a contract if by waiting till time $\bar{\tau}$, P would get a student of expected ability no higher than the expected ability of S given the information available at time τ ; S , in expectation, would get a position of desirability no higher than that of P ; and at least one of these two inequalities is strict.

students are more risk-averse than schools, unraveling may occur even if the equilibrium amount of information is disclosed.

5 Discussion

Our results hinge on the assumption that the supply of jobs of various desirability is fixed exogenously. However, as long as there are rents associated with at least some positions, the model remains relevant. Suppose the desirability of some jobs is fixed exogenously (just like in the case of clerkships, where wages are fixed and are not adjusted depending on the ability of the successful applicant). Wages (and thus desirability) of other positions are chosen by employers in a profit-maximizing manner—we can call these positions competitive market jobs. Productivity of workers in competitive market jobs is proportional to workers' ability level. In this case the distribution of desirability of jobs would adjust in response to any transcript structure so that the average ability of workers at any competitive market job is proportional to the wage. (If all jobs were competitive market jobs, any transcript structure would be an equilibrium outcome and position desirability would be linear in ability.) However, if some jobs are associated with rents, the desirability may increase faster than linear in ability. In other words, our model remains relevant for jobs at the right tail of desirability distribution.²⁴

²⁴Note that in this model the payoff to employers that offer jobs with rent is not specified, all that matters is that these employers prefer to hire candidates with the highest available expected ability.

Notice also that even if schools compete for students, our results remain relevant. We may observe noisy transcripts even if each student knows his ability before starting school and if schools can commit to fully informative transcripts. Of course, in the absence of transfers, if all students know their ability before entering schools, pooling is no longer an equilibrium. If, however, transfers are allowed, we may observe pooling, because less able students are willing to pay to be pooled with more able students. In fact, it can be shown in a formal model that the equilibrium desirability mapping $Q(\cdot)$ remains unchanged if students can negotiate who pools with whom and what the transfers are. (Merit based fellowships might serve a role of transfer payments.)

Another assumption that we make is that a student's ability is exogenously fixed. If learning entails costly effort, noisy transcripts reduce the effort of at least some students (Becker 1982). However, the efficiency loss may be small, because loss in ability is partially compensated by saved effort. At least in theory, grade inflation does not necessarily reduce efficiency. If signaling high ability is merely a ticket to high-rent jobs, then noisy transcripts may be welfare improving. Our present model is not rich enough to analyze efficiency implications of limited information disclosure, and extending it is an interesting area for future research.

Our paper only considers the compression of grades; we do not explain why grades on average go up. Ottaviani and Squintani (2002) offer a model in which grades are inflated because some players do not behave strategically.

Chan, Li, and Suen (2002) assume that employers do not know the underlying distributions of student abilities, and thus can get fooled by schools inflating their grades.

It is often said that “something should be done about grade inflation.” Our results, however, show that unraveling may occur if transcripts at the time of graduation are too informative. In some markets early contracting is a fact of life. For instance, the market for federal judicial law clerks in the US courts has unraveled (Avery, Jolls, Posner, and Roth 2001). Many market participants view unraveling as a serious problem.²⁵ In light of our model we believe that unraveling may occur in that market in part because law schools can not control the amount of information released about students abilities. Indeed, achievements like being an editor of a law review speak volumes about a student’s ability, yet this information is not a part of the transcript and cannot be suppressed by a school.

References

- [1] Avery, Christopher, Andrew Fairbanks, and Richard Zeckhauser (2001), What Worms for the Early Bird? Early Admissions at Selective Colleges, mimeo, Harvard University.

²⁵Avery et. al. give many colorful quotes from judges and students, such as “The unseemly haste to hire law clerks is a disgrace to the federal bench” and “Some judges scrapped decorum and even bare civility. One federal district court judge asked a student to sneak into his office on a Sunday in January, through the service entrance. His court had agreed not to conduct early interviews, he explained, and he wanted to cheat in secret.”

- [2] Avery, Christopher, Christine Jolls, Richard A. Posner, and Alvin E. Roth (2001), The Market for Federal Judicial Law Clerks, University of Chicago Law Review, 68, 3, Summer, 2001, 793-902.
- [3] Becker, William E., Jr. (1982), "The Educational Process and Student Achievement Given Uncertainty in Measurement," The American Economic Review, Vol. 72, No. 1. (Mar., 1982), pp. 229-236.
- [4] Bergemann, Dirk and Martin Pesendorfer (2001), Information Structures in Optimal Auctions, mimeo, Yale University.
- [5] Chan, William, Hao Li, and Wing Suen (2002), Grade Inflation, mimeo, University of Toronto.
- [6] Crawford, Vincent P. and Joel Sobel (1982). Strategic Information Transmission, Econometrica, 50, 6. (Nov., 1982), pp. 1431-1451.
- [7] Haruvy, Ernan, Alvin E. Roth, and M. Utku Unver (2001), The Dynamics of Law Clerk Matching: An Experimental and Computational Investigation of Proposals for Reform of the Market, mimeo, Harvard University.
- [8] Li, Hao and Sherwin Rosen (1998), Unraveling in Matching Markets, American Economic Review, 88(3), June 1998, pp.371-387.
- [9] Li, Hao and Wing Suen (2000), Risk Sharing, Sorting, and Early Contracting, Journal of Political Economy, 108(5), pp. 1058-91

- [10] Ottaviani, Marco, and Francesco Squintani (2002), Non-Fully Strategic Information Transmission, mimeo, University of Rochester.
- [11] Roth, A.E. and X. Xing, Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions, *American Economic Review*, 84, September, 1994, 992-1044.
- [12] Suen, Wing (2000), A Competitive Theory of Equilibrium and Disequilibrium Unravelling in Two-Sided Matching, *RAND Journal of Economics*, 31(1), pp. 101-120.

A Proof of Lemma 1

Suppose in equilibrium students of expected ability \hat{a} get jobs of desirabilities from q_1 to q_2 , $q_1 < q_2$.²⁶ Let \hat{q} be the average desirability that students of expected ability \hat{a} get. $q_1 < \hat{q} < q_2$. Since there is a positive mass of students of expected ability \hat{a} , there must be at least one school producing a positive mass of such students. This school has to include some students of lower and some students of higher ability in this mass. Thus, it can select a small subset from the mass (say, ϵ -share of the mass) such that its expected ability is $\hat{a} - \delta$, where δ is also small. Then the remaining mass has expected ability higher than \hat{a} , and therefore all students there get positions of desirability q_2 or higher. For sufficiently small ϵ and δ , the net change in average desirability is positive, i.e. the school was able to improve upon its equilibrium transcript structure—contradiction.

B Proof of Lemma 5

It is clear that $\hat{a}_L \geq a_L$, since it is impossible to produce students of expected ability lower than the lowest true ability.

Suppose $\hat{a}_L > a_L$. Take a school that has students of true ability a_L (i.e. a positive mass of students of abilities $(a_L, a_L + \epsilon)$ for any positive ϵ). Since the school does not produce any students of ability below \hat{a}_L , it has

²⁶It does not matter for the proof whether this is an open or closed interval. What's important is that all students of expected ability higher than \hat{a} get a job of desirability q_2 or higher.

to “bundle” students in the interval $(a_L, a_L + \epsilon)$ with higher ability students $(0 < \epsilon < \hat{a}_L - a_L)$. But then, since $Q(\hat{a})$ is increasing and convex, the school would increase the average desirability of placements of its students by “unbundling” these low ability students—contradiction.

C Proof of Theorem 1

The proof relies on lemmas of Section 3.2.

Suppose $Q_1 \neq Q_2$. From Lemma 5 we know that they are defined on intervals starting at a_L , and $Q_1(a_L) = Q_2(a_L)$. Let a_D be the first point where these functions begin to differ, i.e. $a_D = \sup\{a \mid \forall a' < a, a' \geq a_L \ Q_1(a') = Q_2(a')\}$, $a_D \geq a_L$. Notice that at least one of these equilibria (without loss of generality, the first one) has to be unfair at a_D (otherwise they would both be fair and thus identical), and so $Q_1(a)$ has to be unfair (and thus linear) on some interval (a_D, x) (If no such x exists, there must exist a decreasing infinite sequence $x_n \rightarrow a_D$ such that Q_1 is fair at x_n for every n . By assumption, Q_T does not switch from convexity to concavity infinitely often, and so for some n' Q_T has to be convex or concave on $[a_D, x_{n'}]$. If it is convex, then mixing is not profitable and the equilibrium has to be fair at a_D ; if it is concave, then Q_1 can not be convex—it equals Q_T at $x_{n'}$, $x_{n'+1}$, and $x_{n'+2}$ —contradiction). Let a_1 be the largest x such that $Q_1(a)$ is linear on (a_D, x) . Take point b such that $Q_2(b) = Q_1(a_1)$.

There are three cases.

- Case 1. $b < a_1$.
- Case 2. $b > a_1$, Q_2 is linear at a_D . Then, by convexity, its slope has to be less than that of Q_1 . Take point a'_1 —the largest x such that Q_2 is linear on (a_D, x) , and point b' such that $Q_1(b') = Q_2(a'_1)$. Then $b' < a'_1$.
- Case 3: $b > a_1$, Q_2 is not linear at a_D . Then there exists a point a''_1 arbitrarily close to a_D at which Q_2 is fair and is less than Q_1 . Take b'' such that $Q_1(b'') = Q_2(a''_1)$. $b'' < a''_1$.

The common feature of all three cases is that we were able to find a point at which ability is truthfully revealed in one equilibrium, and the corresponding wage is higher at another. We now get a contradiction for case 1; for other cases the proof is analogous.

Let $Q^* = Q_1(a_1) = Q_2(b)$. By construction, in the first equilibrium every position below Q^* is filled by someone of higher expected ability than in the second equilibrium. Thus, the total ability of people working at positions below Q^* is higher in the first equilibrium. But let us now look closely at these two populations. By Lemma 4, in the first equilibrium this population is just a group of least able students, since the equilibrium is fair at a_1 . In the second equilibrium it is some mix of the least able students and more able students. But since the total masses of two populations are equal (they are equal to the share of positions below Q^*), the second one has average ability at least as high as the first one—contradiction.

D Proof of Theorem 3

Suppose transcript structures form an equilibrium, and there is no unraveling. We then show that no student has an incentive to deviate, i.e. to sign a contract earlier than $\bar{\tau}$. Consider an arbitrary school i . Let the interval of expected abilities of students at school i at time $\bar{\tau}$ be $[a_i, b_i]$.²⁷ By the law of iterated expectations, no student at school i can have expected ability outside of this interval at any time $\tau \leq \bar{\tau}$. Take any time $\tau < \bar{\tau}$ and any student from school i who has expected ability \hat{a}_τ inside the interval at time τ . If he signs now, the best position he can get is of desirability $Q(\hat{a}_\tau)$. If he waits until time $\bar{\tau}$, the expected desirability of position he gets is $E[Q(\hat{a}_{\bar{\tau}}|\hat{a}_\tau)]$. By assumption, at time $\bar{\tau}$ the school produces a positive density of students on an interval, and transcript structures form an equilibrium—thus, $Q(\hat{a}_{\bar{\tau}})$ is convex on the interval. $E[\hat{a}_{\bar{\tau}}] = \hat{a}_\tau$, and so $E[Q(\hat{a}_{\bar{\tau}}|\hat{a}_\tau)] \geq Q(\hat{a}_\tau)$, and the student does not have an incentive to deviate—this is the same logic as in the proof of Theorem 2.

Now suppose transcript structures do not form an equilibrium, and there is information that a school (say, school i) would like to suppress. This means that by removing this information, and mixing students who are differentiated by it, the school would achieve a higher payoff. But this implies that there exist three points $\varphi_1 < \varphi_2 < \varphi_3$ inside the interval $[a_i, b_i]$ and weight α such that $\varphi_2 = \alpha\varphi_1 + (1 - \alpha)\varphi_3$ and $Q(\varphi_2) > \alpha Q(\varphi_1) + (1 - \alpha)Q(\varphi_3)$.

²⁷Proof is the same if the interval is open at one or both of the ends.

But then there exists some point φ in the open interval (a_i, b_i) such that $Q''(\varphi) < 0$ (otherwise the function would be convex).

Since φ is in the open interval (a_i, b_i) , there exists some $\tau_1 < \bar{\tau}$ such that a positive mass of expected abilities is produced in a small ϵ -neighborhood of φ for any $\tau \in [\tau_1, \bar{\tau}]$. By assumption, $\hat{Q}(a, \tau) = E[Q(\hat{a}_\tau) | \hat{a}_\tau]$ is twice continuously differentiable; also, $\hat{Q}(a, \bar{\tau}) = Q(a)$. Therefore, there exists $\tau_2 < \bar{\tau}$, $\tau_2 \geq \tau_1$ such that $\hat{Q}(a, \tau)''_{aa}(\varphi, \tau) < 0$ for all $\tau \in [\tau_2, \bar{\tau}]$. Finally, there exists $\tau_3 < \bar{\tau}$, $\tau_3 \geq \tau_2$ such that $\sigma(\varphi, \tau) > 0$ for all $\tau \in [\tau_3, \bar{\tau}]$.

By construction, \hat{Q} is a martingale, and therefore $E[d\hat{Q}(a, \tau)] = 0$. By Ito's lemma, $0 = E[d\hat{Q}(a, \tau)] = \frac{1}{2}\sigma^2\hat{Q}''_{aa} + \hat{Q}'_\tau$. For $\tau \in [\tau_3, \bar{\tau}]$, $\frac{1}{2}\sigma^2\hat{Q}''_{aa}(\varphi, \tau) < 0$, and so $\hat{Q}'_\tau(\varphi, \tau) > 0$. But this implies that $Q(\varphi) = \hat{Q}(\varphi, \bar{\tau}) > \hat{Q}(\varphi, \tau_3)$, and so at time τ_3 a student of expected ability φ strictly prefers unraveling and immediately matching with a position of desirability $Q(\varphi)$ to waiting until time $\bar{\tau}$ and getting, in expectation, $\hat{Q}(\varphi, \tau_3)$, while the employer is indifferent.

E The Existence and Connectedness of a Symmetric Equilibrium in a Symmetric Market

Theorem E.1 *If all schools have identical distributions of student abilities, there exists a symmetric equilibrium in pure strategies. This equilibrium is connected.*

Proof. We first show the existence of a symmetric equilibrium in pure

strategies. Let S be the set of a school's strategies. Let $B(s)$ be the best response correspondence—the set of best responses for a school given that all other schools play s . We need to show that $B(s)$ has a fixed point.

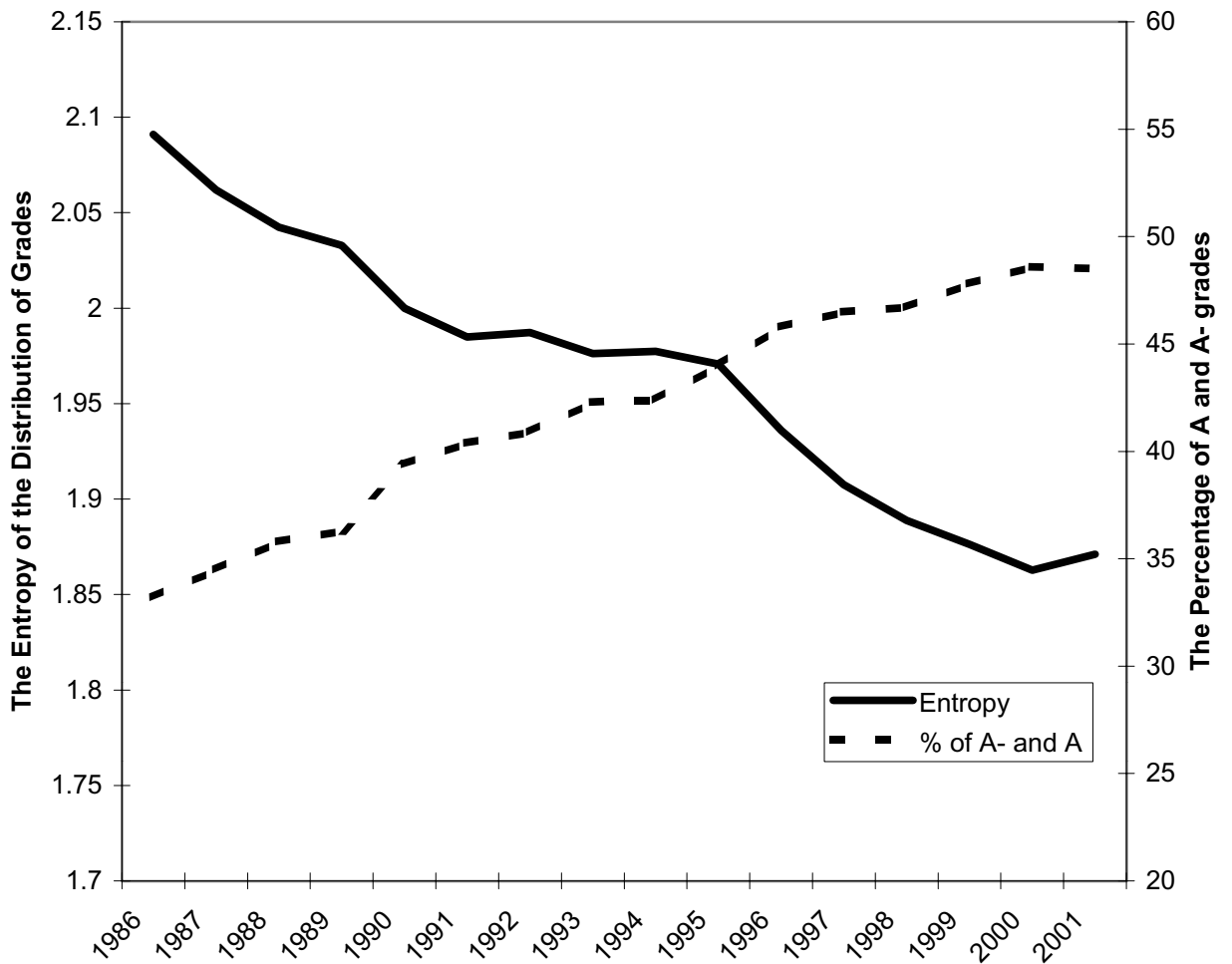
Let S_n be the set of strategies that generate a finite number of transcripts $\{t_i\}$, $i \in \{1, 2, \dots, 2^n - 1\}$, such that transcript $t_{2^{n-1}}$ corresponds to the average true ability in the population, $t_{2^{n-1}+2^{n-2}}$ corresponds to the expected ability of a better-than-average student, $t_{2^{n-1}-2^{n-2}}$ corresponds to the expected ability of a worse-than-average student, and so on (some transcripts may be received by a zero mass of students). S_n is not empty for $n \geq 1$ because it contains a strategy that assigns the same transcript to all students.

S_n is the set of distributions (on the set of $2^n - 1$ points) that second-order stochastically dominate the underlying distribution of student abilities. S_n is convex (if each of two distributions dominates F , their affine combinations do too, and they are also concentrated on the set of $2^n - 1$ points), compact, and the payoff function is continuous on S_n (Each element in S_n is just a vector of $2^n - 1$ positive numbers adding up to 1, and so we can use the induced metric from R^{2^n-1}). Thus, by Kakutani's Fixed Point Theorem there exists s_n^* such that $s_n^* \in B(s_n^*)$, where for a particular n best response correspondence $B(\cdot)$ is restricted to strategies from S_n .

Take the sequence $\{s_n^*\}$ for $n \rightarrow \infty$. It has a limit point in S . This limit point, s , is a best response to itself, and thus corresponds to a symmetric equilibrium.

Let us show that this equilibrium is connected. Suppose it is not. This implies that there is an interval (a, b) such that no school produces students of ability in this interval, but each (since the equilibrium is symmetric) produces positive masses of students on both sides of the interval, i.e. for any open interval containing a or b . But then the school can increase its payoff by mixing some students of ability slightly below a and some students of ability slightly above b so that the expected ability in this mix equals a . ■

Figure 1: The Informativeness of Grades at Harvard



We use entropy as a proxy for the informativeness of grades. It is equal to $(-\sum_i s_i \ln(s_i))$, where i ranges from the worst grade to the best, and s_i is the share of grade i among all grades. If all students receive the same grade, no information is revealed and entropy is minimized. If an informative transcript structure is modified in a way that reduces the amount of information (e.g. students who had a C or a D can no longer be distinguished), entropy goes down.

Figure 2: Probability Density Function of Expected Ability

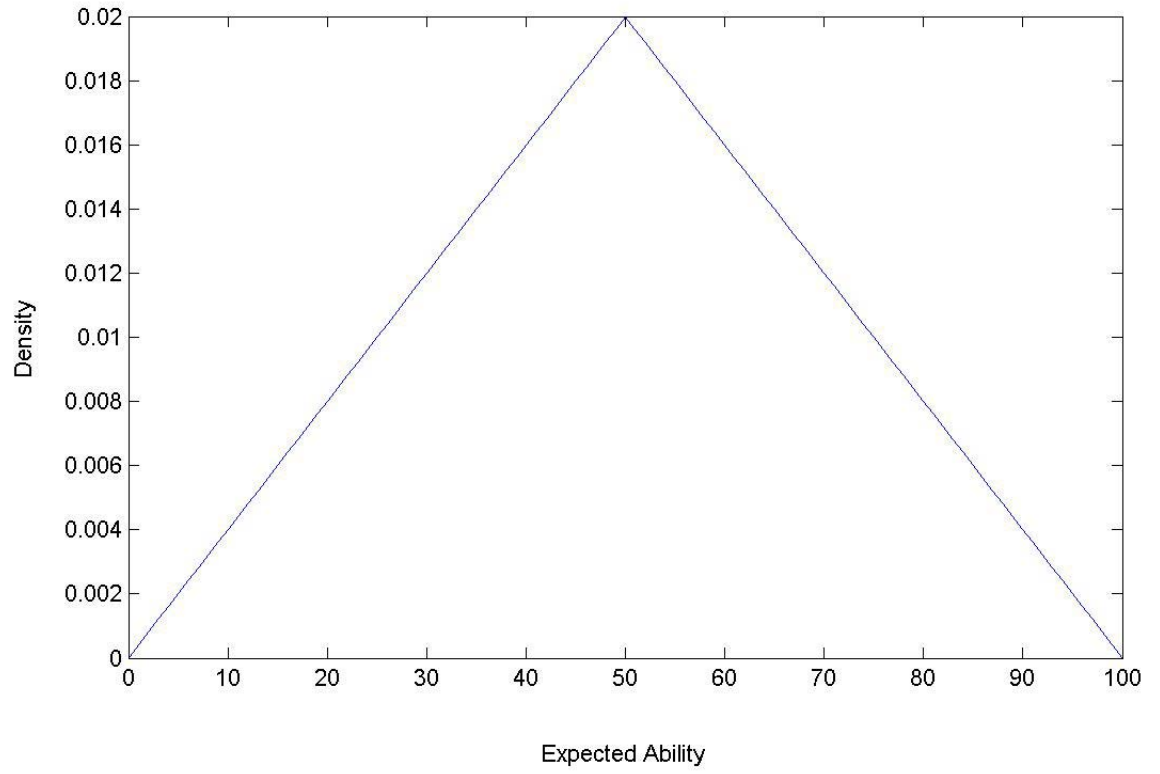


Figure 3: Desirability Mapping

