American Economic Journal: Microeconomics 2 (May 2010): 34–63
http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.2.34

Information Disclosure and Unraveling in Matching Markets[†]

By Michael Ostrovsky and Michael Schwarz*

This paper explores information disclosure in matching markets. A school may suppress some information about students in order to improve their average job placement. We consider a setting with many schools, students, and jobs, and show that if early contracting is impossible, the same, "balanced" amount of information is disclosed in essentially all equilibria. When early contracting is allowed and information arrives gradually, if schools disclose the balanced amount of information, students and employers will not find it profitable to contract early. If they disclose more, some students and employers will prefer to sign contracts before all information is revealed. (*JEL* C78, D82, D83)

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

—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

Labor market institutions often suppress some information about job candidates. For example, students at the Stanford Graduate School of Business (GSB) are graded on a curve, resulting in transcripts that very accurately reflect students' performance. These transcripts, however, are not revealed to potential employers:

> the GSB has no policy on grade disclosure; your grades belong to you and it is your right to use them as you wish. Stanford's nondisclosure norm among MBA students, however, has existed for nearly 40 years.[1]

Most top business schools have similar norms. High profile examples from other areas include Yale Law School, where first semester grades are credit/no credit. Stanford Medical School conceals from residency programs a part of the student's record. Massachusetts Institute of Technology official undergraduate transcripts available to graduate schools and potential employers also suppress available

* Ostrovsky: Graduate School of Business, Stanford University, Stanford, CA 94305 (e-mail: ostrovsky@gsb.stanford.edu); Schwarz: Yahoo! Labs, 2397 Shattuck Ave., Suite 204, Berkeley, CA 94704 (e-mail: mschwarz@yahoo-inc.com). We are grateful to Drew Fudenberg, Kate Ho, Ed Lazear, Muriel Niederle, Al Roth, Andrei Shleifer, Larry Summers, Adam Szeidl, and Robert Wilson for comments and suggestions, and to Tomasz Sadzik for research assistance.

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.2.34.

[1] http://www.gsb.stanford.edu/mba/academics/learning_methods.html, accessed March 22, 2007.

information. They contain only full letter grades, while internal transcripts distinguish between such grades as B+ and B−. Nearly 40 percent of high schools do not disclose class rank to colleges, even though some of them maintain it internally and report it when "absolutely necessary."[2] In fact, more than 90 percent of private nonparochial schools do not disclose class rank (David A. Hawkins et al. 2005).

Concealing information need not require deliberate actions on the part of schools. If revealing full information is not in the interests of a school, it can add noise to transcripts by tolerating (or encouraging) grading policies that make grades less informative. Unless the dean clearly communicates expectations about grading standards, professors are likely to have different ideas regarding the appropriate grade for the average performance. Lack of consistent grading standards adds noise to transcripts, because luck of the draw determines the grading standard adopted by an instructor. A school can reduce this sort of noise by reporting an average grade in each class alongside the grade received by a student or by mandating the use of a forced curve in large classes. Inflated grades could also reduce informativeness of transcripts, perhaps unintentionally. For instance, after years of grade inflation, close to 50 percent of grades in undergraduate classes at Harvard College are A and A−, often erasing the differences between the good and the great. Figure 1 suggests that the informativeness of grades at Harvard, as measured by their entropy, has declined in recent years as the percentage of A and A− grades has risen.[3]

All of the practices described above are similar from the employer's perspective. Refusal to reveal part of the student record, inconsistent grading among instructors, or coarse transcripts are all "noise" that reduces the ability of potential employers to correctly judge the ability of students. The examples above suggest that at least some schools are either indifferent regarding how much information to reveal or prefer to conceal some information about the ability of their students; otherwise, they would try to implement policies that minimize the amount of noise in their transcripts.

There is an alternative channel through which information can be suppressed. Each semester before graduation a student's transcript becomes longer and more informative. Even if schools make transcripts as informative as possible, students and employers may choose to contract significantly before graduation, thus leading to incomplete information disclosure. We say that unraveling occurs when the timing of contracting reduces the amount of available information in a dynamic setting. Early action and early decision admission programs at many selective colleges (Christopher Avery, Andrew Fairbanks, and Richard Zeckhauser 2003) are examples of unraveling. These programs 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). The market for law clerks (Avery et al. 2001) is another dramatic example of unraveling. Avery et al. (2001) report

[2] "Schools Avoid Class Ranking, Vexing Colleges," *New York Times*, March 5, 2006.

[3] Some of Harvard's policies actually encourage grade inflation. An instructor who gives an F or a D is asked to write a note explaining the reasons for the poor performance of a student. In contrast, instructors who give many A grades are not asked to explains their reasons.

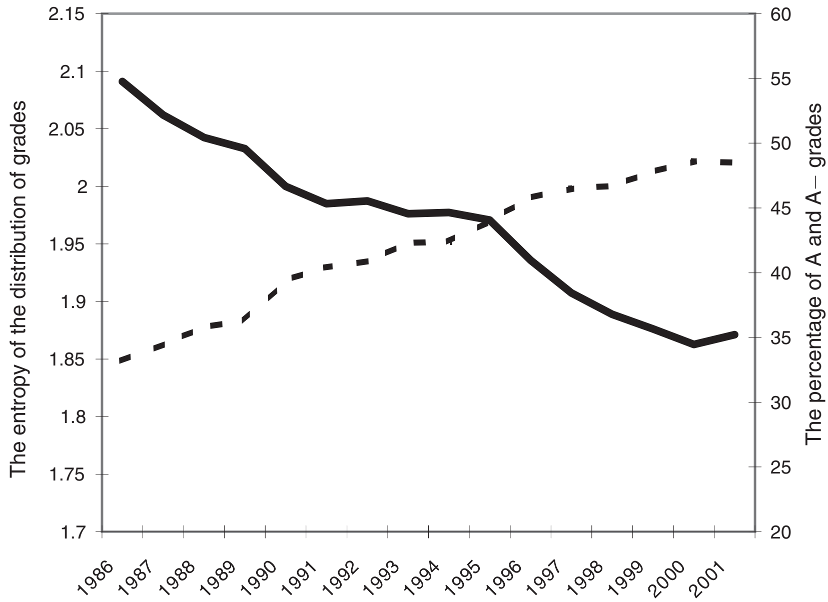

Figure 1. The Informativeness of Grades at Harvard

Notes: The dashed line shows the percentage of A- grades and A grades among all grades at Harvard. The solid line shows the entropy of the distribution of grades. 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 a transcript structure is modified in a way that reduces the amount of information (e.g., students who had a C or a D grade can no longer be distinguished), entropy goes down.

that interviews for clerkship positions are held at the beginning of the second year of three-year law school programs, when only one-third of the students' grades are available. Clearly, a lot of information is withheld. Alvin E. Roth and Xiaoling Xing (1994) describe several other markets in which the timing of transactions has unraveled.

This paper shows that there is a remarkably close connection between the equilibrium (or "balanced") amount of information revealed by schools in the static environment and the incentive for students and employers to unravel in the dynamic environment. If schools disclose the balanced amount of information, students and employers will not find it profitable to contract early. If they disclose more, unraveling will occur.

I. Information Disclosure in a Static Environment

We begin our analysis in the static model. 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.[4] We also assume that wages offered by employers are inflexible, and so the supply of placement slots of a given desirability is exogenously fixed.[5]

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. Outsiders use transcripts 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.

The key feature of our model is that by introducing noise in students' transcripts, a school can change the distribution of desirabilities of positions to which its students are matched in the job market. Consider, for instance, the competition for admission to medical schools. Introducing noise into transcripts may enable a college to increase placement into moderately desirable medical schools at the cost of reducing the number of students placed at top medical schools. The aggregate distribution of positions in the job market does not depend on the transcripts given out by schools, and so the total desirability of placements is constant. However, as we will see in the next section, in a broad range of situations, noise is a necessary feature of transcripts given out in equilibrium.[6]

Consider a population of students. The ability of each student is a real number a in the interval $[a_L, a_H]$. Each student attends one of I schools. The distribution $\lambda_i(\cdot)$ of ability levels at each school i is continuous, exogenous, and commonly known. Without loss of generality, we assume that schools observe the true abilities of their students.[7] Each school decides how much of this information to reveal, i.e., how precise to make its transcripts. A school can make transcripts completely informative, revealing the ability level of each student, or it can make them completely uninformative, or anything in between.

Formally, a school chooses a *transcript structure*, which is a mapping from the abilities of students into *expected abilities* $\hat{a} \in [a_L, a_H]$. This mapping may be stochastic, i.e., for each ability a there can be a probability distribution over the set of expected abilities $\hat{a}$ that a student of ability a can get. However, this mapping has to be statistically correct, in the following sense: the average ability of students "labeled" with expected ability $\hat{a}$ in school i has to equal $\hat{a}$.

[4] The effects of allowing agents in a matching market to invest in their "quality" are analyzed in Harold L. Cole, George J. Mailath, and Andrew Postlewaite (1992, 2001); Michael Peters and Aloysius Siow (2002); Peters (2007); and Ed Hopkins (2010).

[5] One can show that our results remain valid when the wages of some (or all) firms are flexible. Moreover, if the wages of all firms are flexible, then under full information revelation the wage schedule will be convex in ability (see Michael Sattinger 1993, for a survey of the literature on assignment models with flexible wages), and therefore, as we explain in the discussion following Theorem 1, full information revelation by all schools will be an equilibrium outcome.

[6] Several recent papers study strategic disclosure of information in a variety of environments (Steven Matthews and Postlewaite 1985; Masahiro Okuno-Fujiwara, Postlewaite, and Kotaro Suzumura 1990; Alessandro Lizzeri 1999; Archishman Chakraborty and Rick Harbaugh 2007). The distinguishing features of our setup are the general equilibrium approach and the competitive nature of the market.

[7] Suppose 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 an 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."

DEFINITION 1: *A transcript structure is a function $F(\cdot|\cdot)$, where $F(\hat{a}|a)$ is a probability distribution with which a student of ability $a \in [a_L, a_H]$ is mapped to expected ability $\hat{a}$, such that the average ability of students labeled with expected ability $\hat{a}$ is equal to $\hat{a}$.*[8]

Essentially, the definition says that schools give out grades and transcripts to students using some commonly known grading scheme, and then employers can back out each student's expected ability based on his or her transcript, the grading scheme, and the distribution of student abilities in the school. We assume that schools can commit to their transcript structure. This is not a critical assumption.[9] What *is* critical is that employers know the distribution of transcripts given out by a school, as well as the distribution of student abilities there. Employers know the distribution of transcripts if they receive applications from many candidates from a given school. Likewise, the distribution of student abilities in large schools is known to recruiters fairly well, at least if it does not change drastically year-to-year. We rule out the possibility that a school can "fool" employers into thinking that it has better students than it actually does by giving out too many good grades (as in William Chan, Hao Li, and Wing Suen 2007), and focus solely on information compression. This restriction is conceptually similar to the one made by Matthew O. Jackson and Hugo F. Sonnenschein (2007), who show that by linking independent decisions and requiring each agent to report a vector of his realized types in these decisions that mirrors the underlying distribution of types, a mechanism designer can essentially relax incentive constraints as the number of independent decisions becomes large. The key difference between our restriction and theirs is that in our case, while the schools cannot lie "on average," they do have the ability to compress the distribution of reported student abilities, whereas in the case of Jackson and Sonnenschein (2007), the mechanism designer requires the reported distribution of types to coincide with the expected one.

After schools announce transcript structures and announce expected abilities of their students, students and positions are matched. On one side of the market there is a population of students. On the other side of the market there is a set of positions. The desirability of each position, $q \in [q_L, q_H]$, is common knowledge. The distribution $\mu(\cdot)$ of position desirabilities is continuous, exogenous, commonly known, and has positive density on $[q_L, q_H]$. The mass of positions is equal to the mass of students.[10] Students rank positions by desirability, and employers rank students by expected ability.[11] The resulting rankings induce a unique (up to permutations

[8] This definition is very similar to the definition of "information structure" in Dirk Bergemann and Martin Pesendorfer (2007). That paper, however, considers information disclosure in a very different environment (a single-seller, single-object auction), whereas we consider a matching market.

[9] If schools could not commit to their transcript structures, equilibrium information disclosure that we explore in the following section would still remain an equilibrium outcome of the resulting cheap-talk game (see Vincent P. Crawford and Joel Sobel 1982, for a formal analysis of cheap-talk games). Of course, the cheap-talk game has many other equilibria.

[10] This 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.

[11] 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.

of equally desirable positions) assortative stable matching between students and positions.

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.[12]

The following series of examples illustrates the model. In these examples, we discuss equilibrium information disclosure—the concept we formally define in the following section.

Example 1: Student abilities at each school are distributed uniformly on $[0, 100]$, and position desirabilities are also distributed uniformly on $[0, 100]$. If all schools fully reveal student abilities (i.e., set $\hat{a} \equiv a$), the resulting mapping Q from expected abilities to position desirabilities is linear ($Q(\hat{a}) = \hat{a}$), and no school can benefit by deviating. Thus, fully informative transcripts form an equilibrium.

Example 2: Now, suppose that at one-half of all schools, student abilities are distributed uniformly on $[0, 100]$, while the other half has a more able population—student abilities are distributed uniformly on $[50, 100]$. There is a mass 0.5 of students at each type of school. There is also a mass 1 of positions, distributed uniformly on $[0, 100]$, as before. If all schools fully reveal student abilities, the resulting mapping from expected abilities to desirabilities has two linear pieces:

$$Q(\hat{a}) = \begin{cases} \frac{\hat{a}}{2}, & \text{for } \hat{a} \leq 50 \\ \frac{3\hat{a}}{2} - 50, & \text{for } \hat{a} \geq 50. \end{cases}$$

For instance, a student with expected ability 50 is in the twenty-fifth percentile of the student population, and gets a job of the twenty-fifth desirability percentile. Figure 2 illustrates this desirability mapping Q. Note that again, no school can benefit by deviating and suppressing some information. If a "better than average" school mixes some students of different abilities together, it gets exactly the same payoff as without mixing, while if an "average" school mixes students with abilities above 50 and below 50, it gets a strictly lower payoff than without mixing.

Example 3: Finally, suppose that there is an "oversupply" of less able students: at one-half of all schools student abilities are distributed uniformly on $[0, 100]$, while the other half has a *less* able population—student abilities are distributed uniformly on $[0, 50]$. As before, there is a mass 0.5 of students at each type of school and a mass

[12] This can be reconciled with a finite number of schools by using the standard general equilibrium approach; assume that there are I school types and an infinite number of schools of each type. Technically, different schools of the same type could select different transcript structures. However, in any equilibrium in which that could happen, the average transcript structure of schools of a given type would also be an optimal transcript structure for a school of that type to use, and so there exists another equilibrium with the same aggregate properties in which all schools of the same type behave identically.

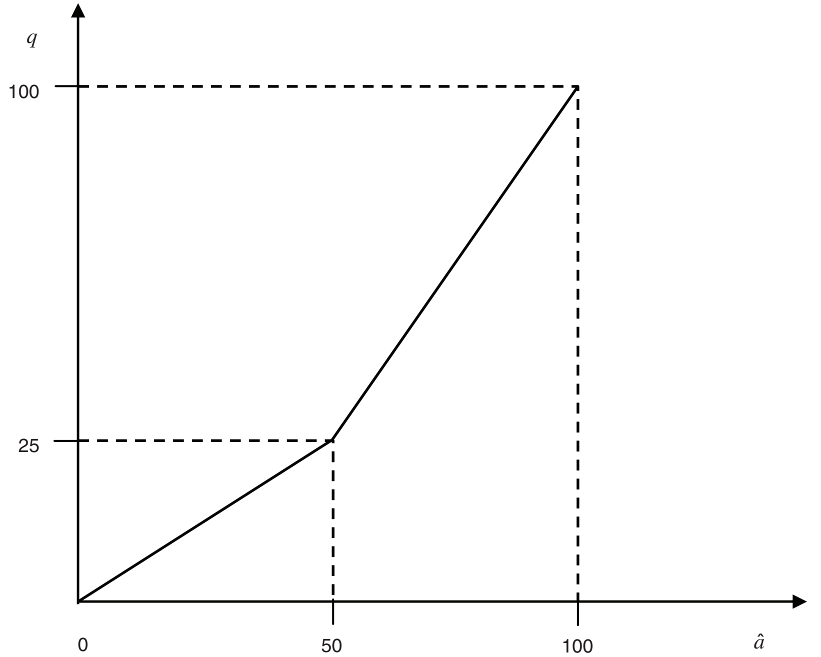

Figure 2. Desirability Mapping Q in Example 2 Under Full Information Revelation

1 of positions, distributed uniformly on [0, 100]. Suppose each school reveals student abilities truthfully. Then the resulting mapping (Figure 3) is

$$Q(\hat{a}) = \begin{cases} \frac{3\hat{a}}{2}, & \text{for } \hat{a} \leq 50 \\ \frac{\hat{a}}{2} + 50, & \text{for } \hat{a} \geq 50. \end{cases}$$

Now, consider a school that contains students of all true abilities from 0 to 100. The average placement obtained by its students is $(75/2 + 175/2)/2 = 62.5$. If the school deviates from full information revelation and, instead, adopts a "no grade disclosure" policy, then every student gets the same transcript, looks the same to employers, and has the expected ability 50. Consequently, the average placement obtained by the school's students increases to 75! Therefore, full revelation is not an equilibrium in this example. What is?

Suppose each "worse than average" school reveals information truthfully, while each "average" school mixes students in such a way that the distribution of expected abilities there is the one plotted in Figure 4, panel B. Then one-third of expected abilities are distributed uniformly on [0, 50], and the remaining two thirds are distributed uniformly on [50, 75].[13] Then the *aggregate* distribution of expected abilities

[13] Note that this distribution second-order stochastically dominates the distribution of true abilities at the "average" school, and therefore there exists a mixing of students generating such distribution of expected abilities.

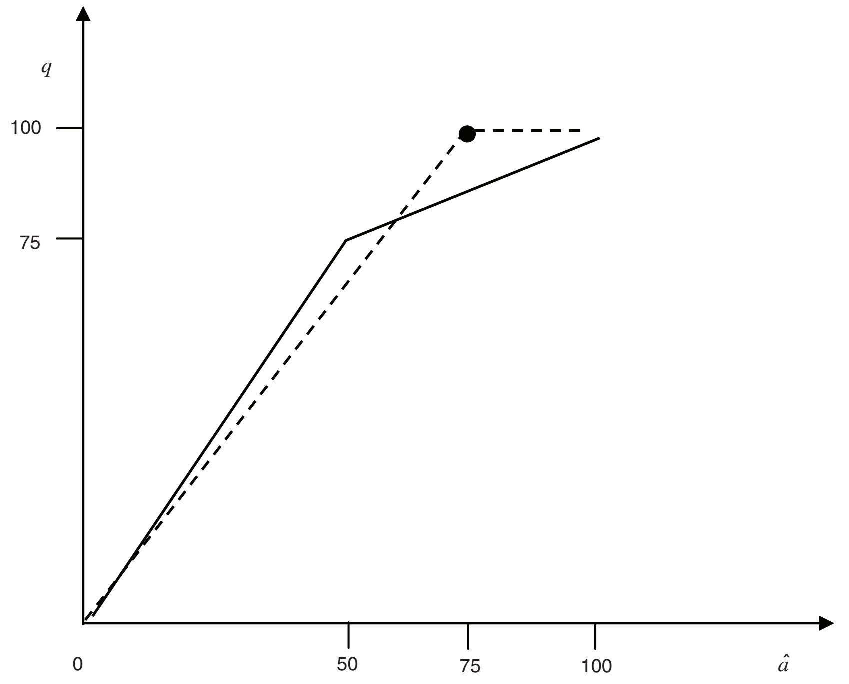

FIGURE 3. DESIRABILITY MAPPINGS IN EXAMPLE 3

Notes: The solid line represents desirability mapping Q arising under full information revelation. The dashed line represents desirability mapping Q' arising in equilibrium.

in the population is uniform on [0, 75], and the corresponding desirability mapping Q' (plotted as the dashed line in Figure 3) is linear:

$$Q'(\hat{a}) = \begin{cases} \frac{4\hat{a}}{3}, & \text{for } \hat{a} \leq 75 \\ 100, & \text{for } \hat{a} \geq 75. \end{cases}$$

This amount of information disclosure is an equilibrium.

II. 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 disclosure of information.

Let $\Phi = (F_1, F_2, \ldots, F_I)$ be a profile of transcript structures, and let G be the aggregate distribution of expected abilities generated by Φ.

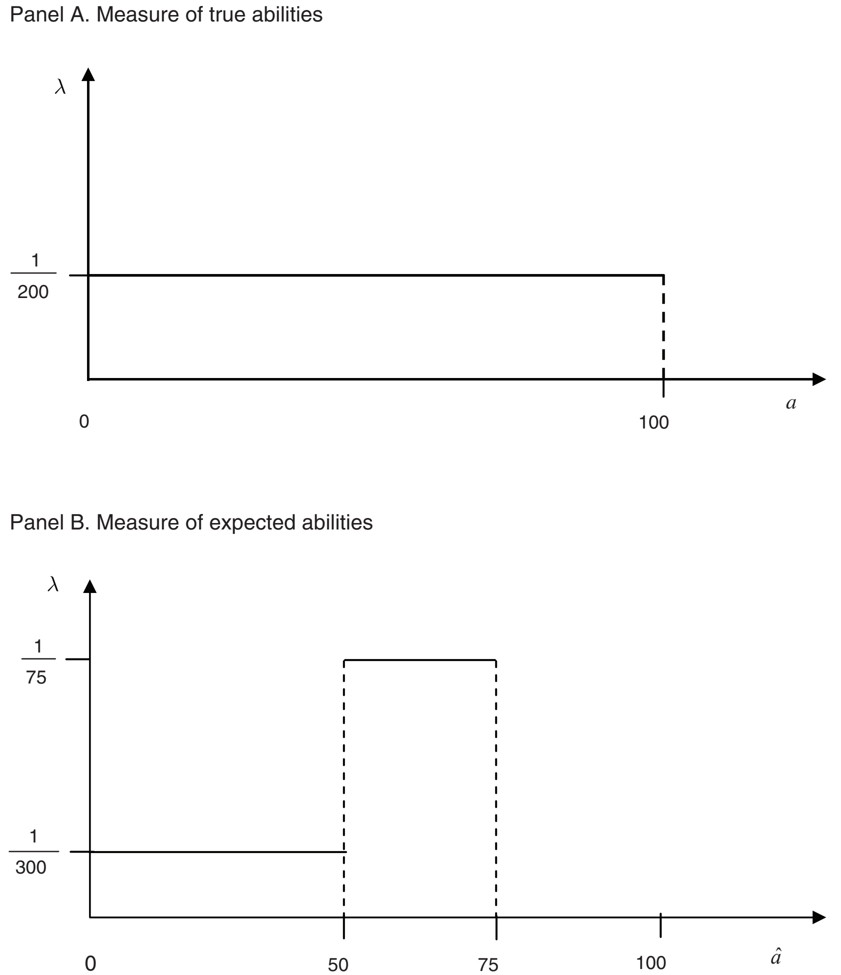

Figure 4. Measures of True and Expected Abilities at an "Average" School in Example 3

DEFINITION 2: *We say that function $Q(\cdot)$ on $[a_L, a_H]$ is the desirability mapping corresponding to profile Φ if, given that schools give out grades in accordance with Φ, the expected desirability of a position matched with a student labeled with expected ability $\hat{a}$ is equal to $Q(\hat{a})$. Formally,*

- $Q(a_L) = q_L$,
- $Q(a_H) = q_H$,
- *for* $\hat{a} \in (q_L, q_H)$,
  - *if* G *is continuous at* $\hat{a}$, *then* $Q(\hat{a}) = \mu^{-1}(G(\hat{a}))$;
  - *if* G *is discontinuous at* $\hat{a}$, *then* $Q(\hat{a}) = \int_{\underline{q}}^{\overline{q}} q\, d\mu(q)/(\mu(\overline{q}) - \mu(\underline{q}))$, *where* $\underline{q} = \mu^{-1}(\lim_{a\to\hat{a}-} G(a))$ *and* $\overline{q} = \mu^{-1}(\lim_{a\to\hat{a}+} G(a))$.

We denote by $Q_T(\cdot)$ the desirability mapping that corresponds to the fully informative profile of transcript structures under which all student abilities are revealed truthfully.

We assume that if all abilities were revealed truthfully, the corresponding desirability mapping $Q_T(\cdot)$ would not switch from convexity to concavity infinitely often. In other words, there exists a finite increasing sequence of ability levels a_i, starting at the lowest and ending at the highest true ability, such that $Q_T(\cdot)$ is convex or concave on each interval $[a_i, a_{i+1}]$.

We now define our solution concept. We say that Φ is an equilibrium profile of transcript structures if each school maximizes the average placement of its students given the desirability mapping $Q(\cdot)$ corresponding to Φ. Formally:

DEFINITION 3: *Take profile Φ, and consider the corresponding aggregate distribution of expected student abilities $G(\cdot)$ and the resulting desirability mapping $Q(\cdot)$. Take any school i, its transcript structure F_i under Φ, and the resulting distribution of expected student abilities in the school, $G_i(\cdot)$. Consider any alternative transcript structure F_i' of school i, and the resulting distribution of expected student abilities, $G'_i(\cdot)$. Profile Φ is an equilibrium profile of transcript structures if for any school i and any alternative transcript structure F_i', the average student placement at school i under the original transcript structure is at least as high as it is under the alternative one, keeping desirability mapping $Q(\cdot)$ fixed:*

$$\int_{a_L}^{a_H} Q(\hat{a})\ dG_i(\hat{a}) \ \geq \ \int_{a_L}^{a_H} Q(\hat{a})\ dG_i'(\hat{a}).$$

Before we can state the main result of this section, we need an additional definition.

DEFINITION 4: *Let $\hat{a}_L$ be the lowest and $\hat{a}_H$ the highest expected ability levels produced in an equilibrium. Then we say that the equilibrium is connected if for every point $\hat{a} \in (\hat{a}_L, \hat{a}_H)$ there exists a school that produces students of all expected abilities in some ε-neighborhood of $\hat{a}$.*

Connectedness is a mild restriction. Indeed, if at least one school gives out some transcripts with the worst and the best possible expected abilities, and everything in between, this restriction is satisfied. Hence, if we observe a school that places students in the entire spectrum of positions, then we can conclude that connectedness holds in the observed equilibrium. Even if such a school does not exist, for connectedness to fail, it would have to be the case that either there is a "hole" in reported ability levels, with some schools producing students of higher ability than a particular interval, some producing students of lower ability than a particular interval, and no school producing students with abilities in that interval; or the boundary in ability levels between some schools is extremely sharp, with one school producing students of abilities arbitrarily close to, but never higher than, some ability level, and another school producing students arbitrarily close to, but never lower than, that ability level. Neither case seems realistic.[14]

Nevertheless, connectedness is a restriction on the solution concept, not on the primitives, and so it is important to address the theoretical question of the existence of connected equilibria. We discuss this question in detail in Appendix C, providing sufficient conditions for the existence of a connected equilibrium, giving an example of a market for which a connected equilibrium does not exist, and describing a general method for checking whether a market has a connected equilibrium.

We are now ready to state and prove the main result of this section. It says that in all connected equilibria, desirability mappings (and, therefore, the aggregate distributions of expected abilities) are the same. In fact, they do not even depend on how students are assigned to schools. Only the aggregate distribution of student abilities and the distribution of position desirabilities matter.[15]

THEOREM 1: *Suppose there is a connected equilibrium with desirability mapping $Q_1(\cdot)$. Suppose students are reshuffled among schools so that the aggregate distribution of student abilities remains the same, but the distributions of abilities within schools possibly change, and suppose there is a new connected equilibrium with desirability mapping $Q_2(\cdot)$. Then for any $\hat{a}$, $Q_1(\hat{a}) = Q_2(\hat{a})$, i.e., the desirability mappings coincide. Equivalently, the aggregate distribution of expected abilities in any connected equilibrium is uniquely determined by the distribution of position desirabilities and the aggregate distribution of true abilities, and does not depend on how these abilities are divided among schools.*

The proof of Theorem 1 proceeds in several steps. First, we show that in any equilibrium, desirability mapping Q is an invertible, monotonically increasing, continuous function, i.e., no positive mass of students receives the same expected ability. Next, we show that in any connected equilibrium, the desirability mapping must be convex. Otherwise, as in Example 3, at least one school will be able to improve its payoff by mixing some students together. On the other hand, if a school does mix students on some interval, the desirability schedule there cannot be strictly convex (and therefore has to be linear). Otherwise, the school would be better off by not mixing the students. Also, we show that the lowest expected ability produced in equilibrium has to be the same as the lowest true ability. The final, and most involved part of the proof, shows that there can only exist one desirability mapping satisfying the above properties for a given pair of distributions of desirabilities and true abilities. This part relies on the assumption that Q_T does not switch from convexity to concavity infinitely often on $[a_L, a_H]$ and proceeds by induction on the number of its inflection points. Along the way, the proof shows how to construct the unique equilibrium desirability mapping and describes what happens in various

[14] In Appendix C, we discuss how a force outside of our model (arbitrage) would tend to eliminate nonconnected, nonconvex equilibria, which may explain why such equilibria are not observed in practice.

[15] Of course, if there existed only one school, and all students went there, this school would be indifferent between all possible amounts of information disclosure. This situation, however, would violate our assumption of price-taking behavior by the schools.

special cases. For instance, if Q_T is convex, then $Q \equiv Q_T$. If Q_T is concave, then Q is linear on $[a_L, \hat{a}_H]$ for some $\hat{a}_H$, and no students have expected abilities above $\hat{a}_H$ in equilibrium. If Q_T is S-shaped, with inflection point $\hat{a}_i$, then in equilibrium there will be "information compression at the top"—up to some level $\hat{a}_* \leq \hat{a}_i$, abilities will be revealed truthfully, and so Q and Q_T will coincide. Above $\hat{a}_*$, students of different abilities will be mixed together, and Q will be linear. Appendix A makes all of these statements formal and gives the full proof of Theorem 1.

Hence, the same amount of information is disclosed in all connected equilibria. We will call this the *balanced amount of information*—the amount that is disclosed in equilibrium when schools can release as little or as much information about their students as they want. Before proceeding further, we give a definition that makes the words "amount of information disclosure" precise. Note that if schools introduce *more* noise in their grades, the resulting distribution of expected abilities gets *compressed*, and its variance *decreases*. This leads to a natural partial ordering on the set of profiles of transcript structures.

DEFINITION 5: *Profile of transcript structures Φ is more informative than profile of transcript structures Φ' if distribution G of expected abilities generated by Φ is second-order stochastically dominated by distribution G' of expected abilities generated by Φ'.*

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.

The last result of this section is a corollary of Theorem 1. It says that if truthful revelation of abilities is an equilibrium (i.e., Q_T is convex), then there are no other connected equilibria.

COROLLARY 1: *Suppose there are multiple connected equilibria in a market, and one of them is fully informative. Then all other equilibria are also fully informative, and therefore the same.*

PROOF:

By Theorem 1, the desirability mappings of all these 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 desirability mapping and the distribution of position desirability). But the fully informative equilibrium is strictly more informative than any other one, and so the other equilibria also have to be fully informative, and therefore the same.

We conclude this section with comments on efficiency implications of information suppression. Our assumptions are insufficient to make unambiguous inferences about efficiency. Indeed, if higher student ability and higher position desirability are complements, then positive assortative matching is efficient, and therefore noisy transcripts will lead to a less efficient allocation of talent than fully informative ones.

If, however, they are substitutes, then negative assortative matching is efficient, and suppressing information will, in fact, lead to a more efficient allocation.

For instance, consider Example 3 from Section I. Suppose the value of interaction between the student of ability a and the position of desirability q is equal to qa, i.e., student ability and position desirability are complements. If schools and students are matched completely randomly, then the total social welfare from the interactions is equal to $50 \times 37.5 = 1{,}875$ (the average position desirability is 50, and the average student ability in the population is 37.5). If schools and students are matched efficiently, i.e., better students are matched to better jobs, then the resulting matching has the desirability mapping shown by the solid line in Figure 3, and, integrating along the vertical axis, the total welfare is equal to

$$\int_0^{100} a(q)q\, d\mu(q) = \int_0^{75} \frac{2q}{3}\, q\, \frac{1}{100}\, dq + \int_{75}^{100} (50 + 2(q - 75))q\, \frac{1}{100}\, dq = 2{,}604.17.$$

Finally, in equilibrium, the desirability mapping is given by the dashed line in Figure 3, and the total welfare is equal to $(1/100) \int_0^{100} (3/4)q^2 dq = 2{,}500$. Thus, efficiency losses due to information suppression in this example are approximately 14 percent (as measured by the total surplus in equilibrium relative to the random assignment versus the socially optimal surplus relative to the random assignment).

Another dimension potentially important for evaluating efficiency is investment in human capital. In our model, a student's ability is exogenously fixed. If learning entails costly effort, noisy transcripts reduce the effort of at least some students (William E. Becker, Jr. 1982). However, the efficiency loss may be small, because the loss in human capital is partially compensated by saved effort. Moreover, if signaling high ability is merely a ticket to high-rent jobs, then noisy transcripts may be welfare-improving.

III. Unraveling

Sections I and II analyze information disclosure in a static framework. In this section, we take the actions of schools as given, but add a time dimension to the model. Students and positions can decide when to sign employment contracts. We show that there is a strong connection between the static concept of "balanced information disclosure" and an inherently dynamic phenomenon that frequently occurs in matching markets—"unraveling," or "early contracting," 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 of high school are available (Avery, Fairbanks, and Zeckhauser 2003); the market for federal judicial law clerks, where judges interview candidates two years prior to the beginning of the clerkships (Avery et al. 2001); and many others (Roth and Xing 1994).

A frequently stated reason for early contracting is insurance. A student may prefer to contract early with a mediocre firm to avoid the possibility of being matched with a really bad firm in case of a negative shock in the future (Li and Sherwin Rosen 1998; Li and Suen 2000; Suen 2000).[16] We consider this explanation in light of our model and establish a close, albeit not obvious, connection between information disclosure and unraveling. To compute the balanced amount of information disclosure, 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 until all information is revealed will lead to the disclosure of more than the balanced amount of information. This is not a coincidence. Theorems 2 and 3 show that if the balanced 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 $Q(\cdot)$, the expected placement that a student will get tomorrow is higher than the placement that he could get today.[17]

It may seem surprising that there is no unraveling under the balanced amount of information disclosure. After all, imagine all positions have similar desirability except for a few that 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 balanced 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 4: Suppose mass 0.8 of position desirabilities is distributed uniformly on [80, 100] ("good jobs"), and mass 0.2 of position desirabilities is distributed uniformly on [0, 80] ("bad jobs"). Suppose also that student abilities in each school are distributed uniformly on [0, 100], and the total mass of students is 1 (Figure 5).

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. The resulting desirability mapping is concave, and so schools can benefit from mixing students (Figure 6). So, what is the balanced amount of information disclosure in this market?

It turns out that the desirability mapping corresponding to the balanced amount of information disclosure in this example is linear on the relevant range:

$$Q'(\hat{a}) = \begin{cases} \dfrac{8}{5}\hat{a}, & \text{for } \hat{a} \leq 62.5 \\ 100, & \text{for } \hat{a} \geq 62.5, \end{cases}$$

[16] Of course, this is not the only possible reason. Unraveling can also occur because of the use of an unstable matching mechanism, exercise of market power, small numbers of participants in a matching market, and other strategic reasons (Roth and Xing 1994; Avery, Fairbanks, and Zeckhauser 2003).

[17] Our arguments rely on the assumption that not only ordinal, but also cardinal preferences of schools and students over positions coincide. Otherwise, unraveling may occur even if the balanced amount of information is disclosed.

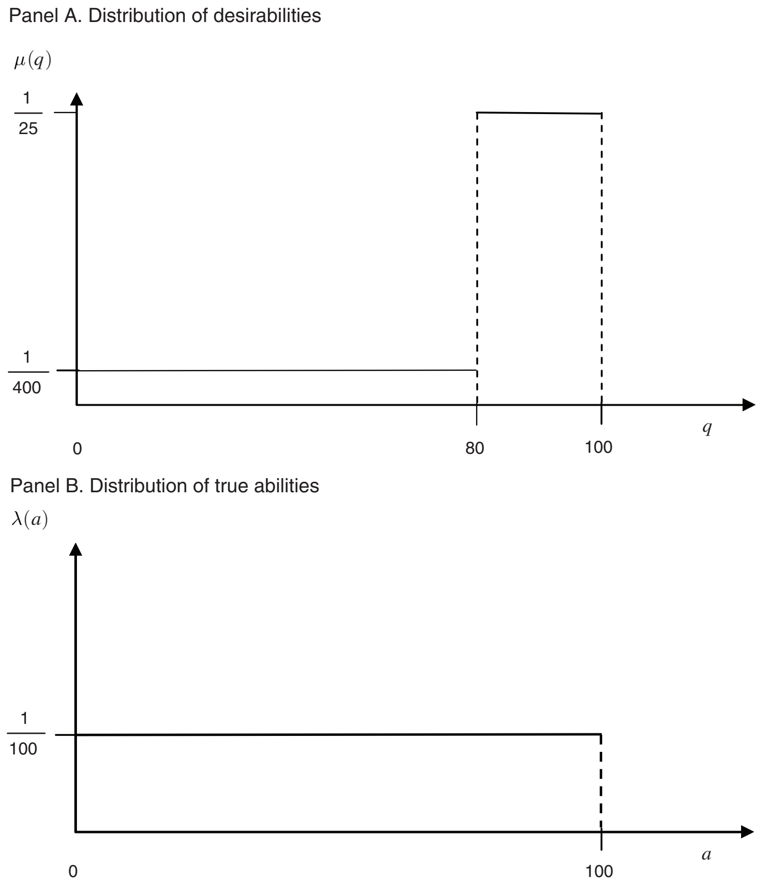

Figure 5. Distributions of Desirabilities and True Abilities in Example 4

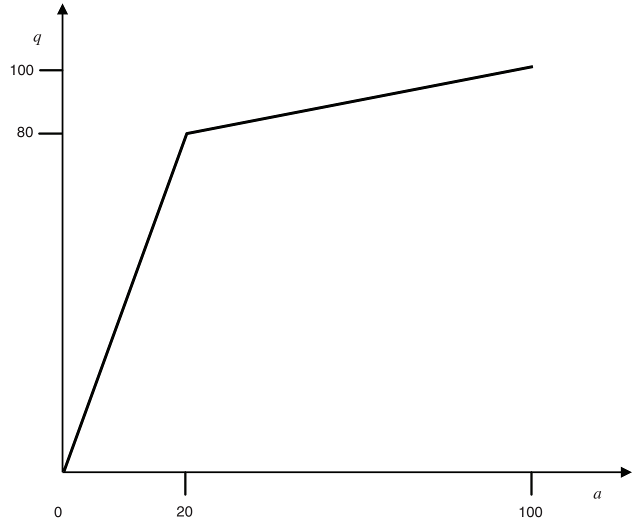

Figure 6. Desirability Mapping Q in Example 4 Under Full Information Revelation

and the corresponding distribution of expected abilities, G, mimics the distribution of position desirabilities:

$$G(\hat{a}) = \begin{cases} \frac{1}{250}\,\hat{a}, & \text{for } \hat{a} \leq 50 \\ \frac{1}{5} + \frac{16}{250}\,(\hat{a} - 50), & \text{for } 50 \leq \hat{a} \leq 62.5. \end{cases}$$

Of course, distribution G has the same mean as the distribution of "true" abilities (uniform on $[0, 100]$), and second-order stochastically dominates it, so there exist transcript structures F_i that give rise to distribution G of expected abilities. Figure 7 illustrates the resulting distribution of expected ability and the corresponding desirability mapping.

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 placement cannot be negative as a result of arrival of new information.

In the remainder of the section, we present a simple two-period model where no information is available in period 1, which is very similar to the model of Suen (2000).[18] This similarity brings into focus the fact that the schools' ability to control information undermines the insurance reason for unraveling. We then show that the result becomes much stronger if information arrives gradually. In that case, if more than the balanced amount of information is disclosed, unraveling will occur.

A. *Two-period Model, Balanced Amount of Information*

Suppose students stay in school for two 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 ability.[19] Suppose employers and students can sign binding contracts in either year of study based on the information available at that period.

[18] We should note that Suen's (2000) 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 in which schools cannot fully control information.

[19] 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 cannot occur under equilibrium information disclosure except for a set of measure zero.

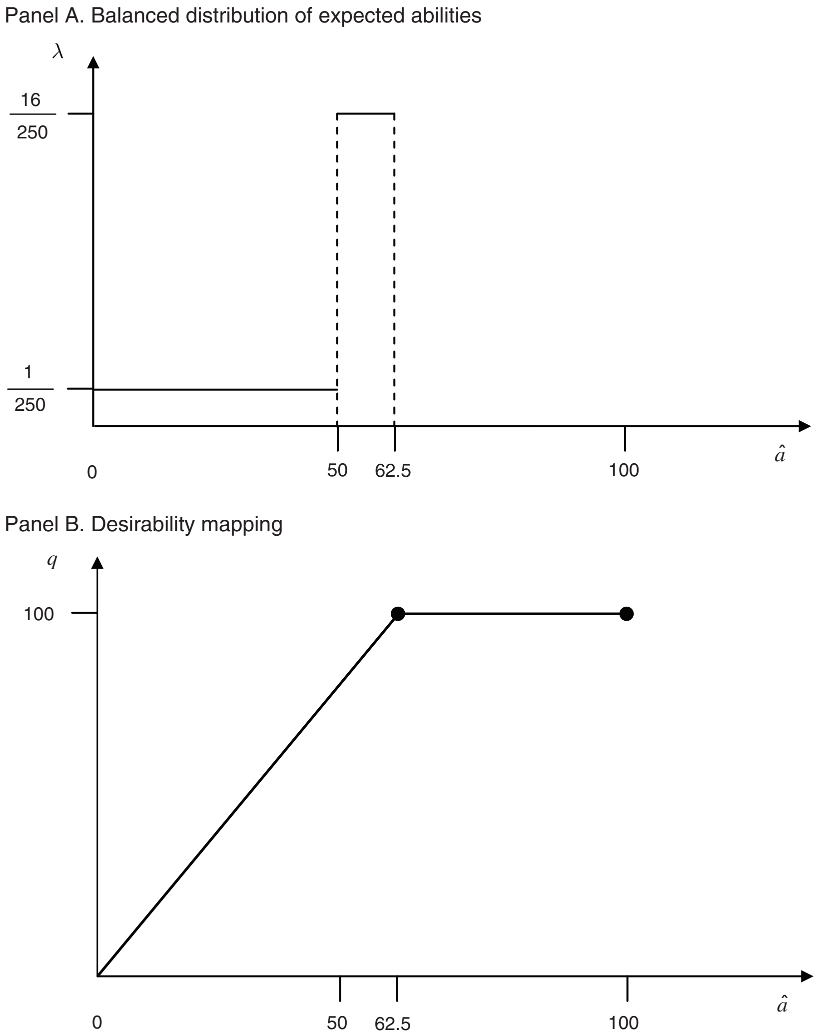

Figure 7. The Balanced Distribution of Expected Abilities and the Corresponding Desirability Mapping Q in Example 4

THEOREM 2: *If, in period* 2, *schools reveal the balanced amount of information, 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, more information about his ability will be revealed. His expected ability will become, say, $\hat{a}$; and he will get a position of desirability $Q(\hat{a})$. By the law of iterated expectations, $E_i[\hat{a}] = \hat{a}_i$. Desirability mapping $Q(\cdot)$ is convex, and therefore $E_i[Q(\hat{a})] \geq Q(\hat{a}_i)$. 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 they cannot benefit from moving early.

B. *Gradual Information Arrival*

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

Students are in school from time $\tau = 0$ until time $\tau = \overline{\tau}$. At time 0 no information about a student is known except for the school i 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 cannot disappear). Namely, at each time τ a potential employer can compute the student's expected ability $\hat{a}_\tau$ based on the school and 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. We assume that $\hat{a}_\tau$ for students in school i follows a diffusion process

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

where diffusion parameter $\sigma_i(\cdot)$ is a bounded continuous function of τ and $\hat{a}_\tau$, such that the process does not leave the interval $[a_L, a_H]$ and for all $\hat{a} \in (a_L, a_H)$, $\sigma_i(\hat{a}, \overline{\tau}) > 0$. We also assume that for some $\tau < \overline{\tau}$, function $\hat{Q}_i(\hat{a}_{\tau'}, \tau') = E[Q(\hat{a}_{\overline{\tau}}) \,|\, \hat{a}_{\tau'}, \tau',\ i]$ is twice continuously differentiable for all $\tau' \in [\tau, \overline{\tau}]$.[20] Whenever expected ability follows such diffusion process, we will say that *information arrives gradually.* Also, we call the amount of information disclosed by schools at the end, i.e., at time $\overline{\tau}$, the *actual amount of information disclosure.*

Each position's desirability is constant and commonly known, and any student-position pair can enter into a binding match at any time. Unraveling occurs if at some time $\tau < \overline{\tau}$ there is a pair, student S and position P, that finds it profitable to sign such a contract.[21]

We now claim that it is an equilibrium for students and firms to sign contracts at time $\overline{\tau}$ without contracting early if the actual amount of information released by schools (i.e., the amount of information disclosed at time $\overline{\tau}$) coincides with the balanced amount of information. If more than the balanced amount of information is disclosed, some students and employers will find it profitable to sign contracts earlier.

THEOREM 3: *Suppose that information about ability of students arrives gradually (see equation* (1)). *If at time $\overline{\tau}$ transcripts contain the balanced amount of information, then it is an equilibrium for all students and positions to wait until time $\overline{\tau}$ to sign contracts. If at time $\overline{\tau}$ transcripts contain more than the balanced amount of*

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

[21] It is profitable for the pair to sign such a contract if by waiting until time $\overline{\tau}$, 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.

information, then some agents are strictly better off not waiting until time $\overline{\tau}$ to sign contracts.

The proof of the first statement of Theorem 3 follows the same intuition as the proof of Theorem 2. If the balanced amount of information is disclosed at time $\overline{\tau}$, desirability mapping $Q(\cdot)$ is convex, making students effectively risk-neutral or risk-seeking, and thus giving them an incentive to wait for additional information.

The proof of the second statement involves several steps. When the actual amount of information disclosed at time $\overline{\tau}$ is between the balanced and the full amounts, we show that at the points of strict convexity of the balanced desirability mapping all three desirability mappings coincide, and students of abilities below and above such points are not mixed together. Thus, any additional information revelation in the actual versus the balanced amounts has to take place in an interval where the balanced desirability mapping is linear. But then at some expected ability level $\hat{a}_*$ in this interval, the actual desirability mapping $Q(\cdot)$ will be locally concave. This implies that at some time τ sufficiently close to $\overline{\tau}$, a student of expected ability $\hat{a}_*$ will strictly prefer immediately signing a contract with a position of desirability $Q(\hat{a}_*)$ to waiting until time $\overline{\tau}$. See Appendix B for the detailed proof.

IV. Conclusion

Information suppression by schools is a widespread phenomenon, taking many forms from nondisclosure policies to coarse transcripts and inconsistent grading standards. We show that such behavior may be necessary in equilibrium. For many distributions of student abilities and job desirabilities, if all schools revealed full information about their students, then some of them could benefit by giving similar transcripts to students of different abilities, thus increasing the placement into moderately desirable positions and reducing the placement into very desirable and very undesirable ones. We also show that an essentially unique amount of information is disclosed in all equilibria. We call this the balanced amount of information disclosure.

Schools are not the only actors in this market who can suppress information. By signing contracts early, students and employers can forgo the information about the students' performance in the last few semesters. We show that these two seemingly distinct ways of suppressing information are, in fact, closely related. If schools disclose the balanced amount of information, students and employers will not find it profitable to contract early. If they disclose more, unraveling will occur.

The intuition behind this connection is very natural. Under the balanced amount of information disclosure, the mapping from expected student abilities inferred from their transcripts to job desirabilities must be convex; otherwise, a school could "mix" some students together and increase its payoff. Hence, if the balanced amount of information is disclosed at graduation, a student is effectively risk-neutral or even risk-seeking. In expectation, additional information does not hurt him. If, however, schools disclose more than the balanced amount of information, some parts of the ability–desirability mapping become concave, and students in the relevant ability range become effectively risk-averse, thus trying to avoid the arrival of future information by contracting early.

Educational institutions are often criticized for not revealing full and accurate information about their students, either by means of nondisclosure policies[22] or as a result of grade inflation, which can compress grades so that they lose some of their informativeness.[23] Our results show that information suppression may be inevitable. Even if schools reveal full information about their students, some of that information will be suppressed via a different channel—unraveling will occur. Unraveling in various markets and its consequences are documented in Roth and Xing (1994), Avery et al. (2001), and Muriel Niederle and Roth (2003). Avery et al. (2001) give many colorful quotes from judges and law school students who experience the effects of unraveling in the market for federal judicial law clerks, 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."[24] Anyone who claims that more information needs to be disclosed has to keep in mind the "unseemly haste" that may follow.

Appendix A: Proof of Theorem 1

We first show that in equilibrium there is a one-to-one mapping from expected ability to position desirability, i.e., the distribution of *expected* abilities, as well as the corresponding desirability mapping Q are continuous in equilibrium. This implies that Q is an invertible function.

LEMMA A1: *In equilibrium, any two students of the same expected ability $\hat{a}$ obtain equally desirable positions.*

PROOF:

Suppose in equilibrium students of expected ability $\hat{a}$ get jobs of desirabilities from q_1 to q_2, $q_1 < q_2$, i.e., there is a positive mass of students of expected ability $\hat{a}$. 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 ability 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 an 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.

Desirability mapping $Q(\cdot)$ is monotonically increasing. This, however, does not necessarily mean that a student of a higher true ability will get matched to a better

[22] "Campus Confidential," *Business Week*, September 12, 2005.

[23] "The Economics of Grade Inflation," *The Economist*, March 7, 2002.

[24] The quote continues, "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."

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.

We will say that an equilibrium is *fully informative* 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 *fully informative* at desirability q and ability a if $Q(a) = q$, no students with true ability below a get matched with jobs of desirability above q, and no students with true ability above a get matched with jobs of desirability below q. It is straightforward to show that an equilibrium is fully informative (i.e., schools do not suppress any information) if and only if it is fully informative at every position desirability.

Now, suppose a school produces students of expected abilities 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.

Next, if a school does mix students of true abilities b and c, by convexity of desirability mapping $Q(\hat{a})$, this could only be optimal if the desirability mapping is linear on the interval $[b, c]$. Consequently, if $Q(\hat{a})$ is strictly convex at a certain expected ability level a, it is fully informative at $Q(a)$: students with ability above a get positions better than $Q(a)$, and students with ability below a get positions worse than $Q(a)$. Therefore, in that case, $Q(a) = Q_T(a)$ (recall that $Q_T(a)$ is the desirability mapping that would arise if all schools revealed all abilities truthfully).

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 signaling" in a separating equilibrium of a signaling game.

LEMMA A2: *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:

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 $(a_L, a_L + \varepsilon)$ for any positive ε). Since the school does not produce any students of ability below $\hat{a}_L$, it has to "bundle" students in the interval $(a_L, a_L + \varepsilon)$ with higher ability students $(0 < \varepsilon < \hat{a}_L - a_L)$. 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.

We are now ready to prove Theorem 1. The proof proceeds by induction on the number of intervals on which the convexity or concavity of $Q_T(a)$ does not change (and, along the way, shows how to construct the equilibrium desirability mapping). For convenience, we will call such intervals "convexity intervals." Recall that by assumption, $Q_T(a)$ does not switch from convexity to concavity infinitely often, and hence has a finite number of convexity intervals.[25]

Step 1: Suppose $Q_T(a)$ has only one convexity interval.

Step 1: Case "Convex." Suppose $Q_T(a)$ is convex on $[a_L, a_H]$. Then truthful revelation is an equilibrium profile of transcript structures. Suppose there is another equilibrium profile of transcript structures Φ, involving some mixing of students, with desirability mapping $Q(\hat{a})$ on $[a_L, \hat{a}_H]$, where $\hat{a}_H \leq a_H$. Take any point x_1 on $(a_L, \hat{a}_H)$ such that $Q_T(x_1) \neq Q(x_1)$. Equilibrium Φ is not fully informative at x_1, and is therefore linear on some interval containing x_1. Take the largest such interval $[a_1, a_2]$. Equilibrium Φ has to be fully informative at a_1, and therefore $Q_T(a_1) = Q(a_1)$.

With a_2, there are two possibilities.

If $a_2 < \hat{a}_H$ or $a_2 = \hat{a}_H = a_H$, then Φ also has to be fully informative at a_2, with $Q_T(a_2) = Q(a_2)$. But then Q_T is convex, Q is linear on $[a_1, a_2]$, $Q_T(a_1) = Q(a_1)$, $Q_T(a_2) = Q(a_2)$, and $Q_T(x_1) \neq Q(x_1)$ (with $a_1 < x_1 < a_2$), which implies that $Q_T(x_1) < Q(x_1)$, which in turn implies that for all $x \in (a_1, a_2)$, $Q_T(x) < Q(x)$. This, in turn, implies that *every* firm of desirability q strictly between $q_1 = Q_T(a_1) = Q(a_1)$ and $q_2 = Q_T(a_2) = Q(a_2)$ is matched to a better (in expectation) student under truthful revelation than under equilibrium with mixing Φ, which, finally, implies that the *total* ability of students matched to those positions in equilibrium Φ is strictly higher than the total ability of students matched to them under truthful revelation, i.e.,

$$\int_{q_1}^{q_2} \left(a_1 + \frac{q - q_1}{q_2 - q_1}\,(a_2 - a_1) \right) d\mu(q) < \int_{q_1}^{q_2} Q_T^{-1}(q)\; d\mu(q) = \int_{a_1}^{a_2} a\, dG(a).$$

This is impossible, because, by construction, desirability mapping Q is strictly convex at both a_1 and a_2, and so the set of students matched with positions in the range $[q_1, q_2]$ in equilibrium Φ is the same as the set of students matched with those positions under truthful revelation, and so all of the integrals above have to be equal.

If $a_2 = \hat{a}_H < a_H$, then $Q_T(\hat{a}_H) < Q_T(a_H) = Q(\hat{a}_H)$, and by convexity of Q_T and linearity of Q on $[a_1, \hat{a}_H]$, for all $x \in (a_1, \hat{a}_H)$, $Q_T(x) < Q(x)$, and therefore for all $q \in (Q_T(a_1), Q_T(a_H))$, $Q_T^{-1}(q) > Q^{-1}(q)$, which is impossible because equilibrium Φ is fully informative at a_1, and the set of students matched to positions above $Q_T(a_1)$ is the same under Φ and under truthful revelation (in both cases, it is the set of students with abilities above a_1).

Step 1: Case "Concave." Suppose $Q_T(a)$ is concave on $[a_L, a_H]$. In equilibrium, the desirability mapping Q has to be linear on the entire interval $[a_L, \hat{a}_H]$ for some $\hat{a}_H \leq a_H$. Indeed, suppose there is a point, $\hat{a}$, at which Q is not linear. Then it

[25] We do not discuss in detail intervals on which Q_T is linear, and effectively assume its interval-wise strict concavity or convexity. Considering the intervals on which the mapping is linear is not hard conceptually, but would make the proof more cumbersome.

has to be strictly convex (and equilibrium fully informative) at $\hat{a}$. By an argument analogous to that of Case "Convex," this is impossible.

Moreover, there exists only one $\hat{a}_H$ that can arise in equilibrium. It is the unique one that guarantees that the total ability of students assigned to all schools is equal to the total ability of students in the population, i.e., the unique $\hat{a}_H$ such that

$$\int_{q_L}^{q_H} \left(a_L + \frac{q - q_L}{q_H - q_L} \left(\hat{a}_H - a_L \right) \right) d\mu(q) = \int_{a_L}^{a_H} a \, dG(a).$$

Step 2: We are now ready to prove the inductive step. Suppose the theorem is true for all $n < k$, and suppose there are $k > 1$ convexity intervals in Q_T. Take the first one, i.e., the one that begins at a_L and ends at some value b_1. It is now more convenient to consider the two cases in the reverse order.

Step 2: Case "Concave." Suppose Q_T is concave on $[a_L, b_1]$. By an argument analogous to the one above, equilibrium desirability mapping Q has to be linear on interval $[a_L, c_1]$ for some $c_1 > b_1$. Let us find this point c_1 and show that it is uniquely determined. Consider the graph of Q_T on a two-dimensional plane, and take the infinite ray that starts at the point (a_L, q_L) and has a slope of zero. Start rotating this ray around its origin, increasing its slope. Once the ray begins to intersect with the graph of Q_T at points (a_i, q_i) other than the origin, for each of these points (and there is always a finite number of them, at most two per convexity interval) keep checking whether they could potentially be the c_1 we are looking for. Specifically, check whether the total ability of all students of ability below a_i is equal to the hypothetical total ability of students assigned to positions of quality below q_i under the linear desirability mapping implied by the ray, i.e., whether

$$\int_{a_L}^{a_i} a \, dG(a) = \int_{q_L}^{q_i} \left(a_L + \frac{q - q_L}{q_i - q_L} \left(a_i - a_L \right) \right) d\mu(q).$$

As soon as such a point exists, stop, and consider this point (a^*, q^*). If, by coincidence, there are several such points on the ray, consider the one with the largest coordinates. This is Subcase "Partially Linear." If no such point exists for any slope less than or equal to $(q_H - q_L)/(a_H - a_L)$, let $q^* = q^H$, take the unique point (a^*, q^H) such that the total ability of students assigned to positions $[q_L, q_H]$ implied by desirability mapping $q_L + ((q_H - q_L)/(a^* - a_L))(\hat{a} - a_L)$ is equal to the total ability of all students in the population, i.e., $\int_{q_L}^{q_H} \left(a_L + ((q - q_L)/(q_H - q_L))(a^* - a_L) \right) d\mu(q) = \int_{a_L}^{a_H} a \, dG(a)$; this is Subcase "Fully Linear" below.

We now claim that in any connected equilibrium, $c_1 = a^*$ and the desirability mapping on $[a_L, a^*]$ is a straight line between (a_L, q_L) and (a^*, q^*).

Step 2: Case "Concave." Subcase "Partially Linear." Suppose there exists an equilibrium, Φ, for which $c_1 \neq a^*$. Consider students assigned to positions $[q_L, q^*]$

under Φ and under truthful revelation. Under truthful revelation, matching is based on true ability, and so these positions get the worst possible students. Hence, the total ability of these students has to be at most as high as the total ability of students assigned to these positions under Φ. Now, consider desirability mapping Q corresponding to Φ. By construction, the slope of Q at a_L is at least as high as $(q^* - q_L)/(a^* - a_L)$ and $Q(a^*) > q^*$, which implies that for all $q \in [q_L, q^*]$, $Q^{-1}(q) \leq a_L + ((q - q_L)/(q_i - q_L))(a_i - a_L)$, and for a positive mass of positions q from this interval, $Q^{-1}(q) < a_L + ((q - q_L)/(q_i - q_L))\,(a_i - a_L)$. But this leads us to a contradiction, because then the total ability of students assigned to positions $[q_L, q^*]$ under Φ, $\int_{q_L}^{q_i} Q^{-1}(q)\, d\mu(q)$, is strictly less than $\int_{q_L}^{q^*} \left(a_L + ((q - q_L)/(q^* - q_L))\,(a^* - a_L)\right)\, d\mu(q)$, which by construction is equal to $\int_{a_L}^{a_i} a\, dG(a)$, i.e., the total ability of students assigned to these positions under truthful revelation.

To complete the inductive step for this case, it is now sufficient to note that if $a^* = a_H$, then we are done. Otherwise, the equilibrium desirability mapping for expected ability levels above a^* is uniquely determined as the equilibrium desirability mapping of the original economy excluding the students of ability below a^* and positions of desirability below q^*; in this truncated economy, the number of convexity intervals is less than k, satisfying the assumptions of the inductive step.

Step 2: Case "Concave." Subcase "Fully Linear." This substep follows from the same ideas as subcase "Partially Linear" and case "Concave" of Step 1, and is therefore omitted.

Step 2: Case "Convex." Suppose Q_T is convex on $[a_L, b_1]$. Our method for finding the unique equilibrium desirability mapping Q is based on the following observation. Suppose Q and Q_T do not coincide on $[a_L, b_1]$. Then there exists $a \in [a_L, b_1)$ such that

1) $Q(x) = Q_T(x)$ for all $x \in [a_L, a]$,
2) $Q(x)$ is linear on $[a, b_1]$,
3) the slope of $Q(x)$ on $[a, b_1]$ is less than or equal to the right derivative of $Q_T(x)$ at a, and
4) if $a > a_L$, the slope of $Q(x)$ on $[a, b_1]$ is greater than or equal to the left derivative of $Q_T(x)$ at a.

Indeed, suppose for some $x \in (a_L, b_1)$, $Q(x) \neq Q_T(x)$. Then we know that $Q(x)$ has to be linear on some interval around x. Take the largest such interval $[a, b]$. The equilibrium is fully informative at a. It is also fully informative at a_L. If $a = a_L$, statement (1) above follows trivially; otherwise it follows from the convexity of $Q_T(x)$ on $[a_L, a]$ by the argument analogous to Step 1, case "Convex." By a similar argument, b has to be strictly greater than b_1, giving us (2). Statement (4) follows immediately from the convexity of equilibrium desirability mapping Q at any point, including, of course, a. To prove (3), suppose the slope of $Q(x)$ on $[a, b_1]$ is strictly greater than the right derivative of Q_T at a. Then the total ability of students assigned to positions in some small interval $[Q(a), Q(a) + \varepsilon]$ in this equilibrium is strictly lower than the total ability of students assigned to these positions under full information revelation, which is impossible.

Let $\overline{r}$ be the left derivative of Q_T at b_1. The observation above implies that any equilibrium has to be either fully informative on $[a_L, b_1]$ or fully informative up to some ability level $a \geq a_L$ and linear with some slope $r < \overline{r}$ on $[a, b_1]$. Crucially, it also implies that for any slope $r < \overline{r}$, there exists exactly one point $a(r)$ on $[a_L, b_1]$, at which an equilibrium could switch from being fully informative to being linear with slope r. This point $a(r)$ is simply the point at which a line with slope r is tangent to the graph of function Q_T.

We now proceed in essentially the same way as in Step 2, Case "Concave," starting with a ray from (a_L, q_L) and a slope of zero, and gradually increasing the slope, looking first for a "Partially Linear" subcase and then, after the ray crosses the point (a_H, q_H), looking for the "Fully Linear" subcase. There are, however, two differences. First, as we increase the slope, we also gradually move the origin of the ray along the graph of mapping Q_T, keeping the ray tangent to it. Second, the slope may reach $\overline{r}$ before encountering either the "Partially Linear" or "Fully Linear" subcase. If that happens, we know that the equilibrium has to be fully informative on $[a_L, b_1]$, and the rest of the desirability mapping is uniquely pinned down as the equilibrium desirability mapping of the original economy excluding the students of ability below b_1 and positions of desirability below $Q_T(b_1)$. In this truncated economy, the number of convexity intervals is equal to $k - 1$, satisfying the assumptions of the inductive step.

Appendix B: Proof of Theorem 3

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

Now suppose more than the balanced amount of information is disclosed at time $\overline{\tau}$. We first show that the corresponding desirability mapping is not convex. Let F be the distribution of expected abilities under balanced information disclosure, G the distribution of expected abilities actually disclosed at time $\overline{\tau}$, and H the distribution of true abilities. We know that H is more informative than G, which, in turn, is more informative than F. Suppose at some expected ability level a, the desirability mapping corresponding to F is strictly convex. Then, as we have previously explained, in the static equilibrium, schools do not mix students with abilities below a and above a, and so under both F and H, a student with reported expected ability a gets matched with a job of the same desirability d. Moreover, the average (and the total) ability of students matched with positions of desirabilities less than d is the same under F and H.

Consider two arbitrary distributions of expected abilities, β and γ, desirability level δ, and expected ability level α corresponding to δ if expected abilities are distributed according to β. Note that if γ is less informative than β, then the average (or total) ability of students matched with positions of desirability less than δ under γ is a least as large as under β. Moreover, the two are equal only if distribution β restricted to $[a_L, \alpha]$ is a mean-preserving spread of distribution γ restricted to the same interval, i.e., under γ (relative to β), students of expected ability levels below α do not get mixed with students of expected ability levels above α.

But then it has to be the case that under distribution G, which in terms of informativeness is between distributions F and H, students of ability below a do not get mixed with students of ability above a. Therefore, any piece of additional disclosure of information under G versus F has to take the form of a mean-preserving spread of the distribution of expected abilities in a region where the desirability mapping under F is linear. It is easy to see that any amount of additional information in such a region leads to a desirability mapping that is not convex. Hence, there exists some point $\hat{a}_*$ inside that region such that $Q''(\hat{a}_*) < 0$.

Since $\hat{a}_*$ is inside the region of reported abilities, there exists some $\tau_1 < \overline{\tau}$ and school i such that a positive mass of expected abilities is produced by school i in a small ε-neighborhood of $\hat{a}_*$ for any $\tau \in [\tau_1, \overline{\tau}]$. By assumption, $\hat{Q}_i(\hat{a}_\tau, \tau) = E[Q(\hat{a}_{\overline{\tau}}) \,|\, \hat{a}_\tau, \tau, i]$ is twice continuously differentiable; also, $\hat{Q}_i(\hat{a}, \overline{\tau}) = Q(\hat{a})$. Therefore, there exists $\tau_2 < \overline{\tau}$, $\tau_2 \geq \tau_1$ such that $(\partial^2 \hat{Q}_i(\hat{a}_*, \tau))/(\partial \hat{a}^2) < 0$ for all $\tau \in [\tau_2, \overline{\tau}]$. Finally, there exists $\tau_3 < \overline{\tau}$, $\tau_3 \geq \tau_2$ such that diffusion parameter $\sigma_i(\hat{a}_*, \tau)$ is strictly positive for all $\tau \in [\tau_3, \overline{\tau}]$.

By construction, $\hat{Q}_i$ is a martingale, and therefore $E[d\hat{Q}_i(\hat{a}, \tau)] = 0$. By Ito's lemma,

$$0 = E[d\hat{Q}_i(\hat{a}, \tau)] = \frac{1}{2}\sigma_i^2 \frac{\partial^2 \hat{Q}_i(\hat{a}, \tau)}{\partial \hat{a}^2} + \frac{\partial \hat{Q}_i(\hat{a}, \tau)}{\partial \tau}.$$

For $\tau \in [\tau_3, \overline{\tau}]$,

$$\frac{1}{2}\sigma_i^2 \frac{\partial^2 \hat{Q}_i(\hat{a}_*, \tau)}{\partial \hat{a}^2} < 0,$$

and so

$$\frac{\partial \hat{Q}_i(\hat{a}_*, \tau)}{\partial \tau} > 0.$$

But this implies that $Q(\hat{a}_*) = \hat{Q}_i(\hat{a}_*, \overline{\tau}) > \hat{Q}_i(\hat{a}_*, \tau_3)$, and so at time τ_3, a student of expected ability $\hat{a}_*$ strictly prefers unraveling and immediately matching with a position of desirability $Q(\hat{a}_*)$ to waiting until time $\overline{\tau}$ and getting, in expectation, $\hat{Q}_i(\hat{a}_*, \tau_3)$, while the employer is indifferent.

Appendix C: The Connectedness Restriction

In this Appendix, we discuss the connectedness restriction. First, we give a sufficient condition for the existence of a connected equilibrium—all schools are identical. Next, we show that in some markets, a connected equilibrium may not exist, even if for every *true* ability level $a \in (a_L, a_H)$ there exists a school that has students of all *true* abilities in some ε-neighborhood of a. Third, we present a method for checking whether a connected equilibrium exists. We conclude by discussing how arbitrage would tend to eliminate equilibria with nonconvex desirability mappings.

THEOREM C.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 prove 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 correspondence $B(\cdot)$ has a fixed point.

Define S_n as the set of all strategies that generate a finite number of expected abilities $\{\hat{a}_i^n\}$, $i \in \{1, 2, \ldots, 2^n - 1\}$, such that expected ability $\hat{a}_{2^{n-1}}^n$ corresponds to the average true ability in the population, $\hat{a}_{2^{n-1}+2^{n-2}}^n$ corresponds to the expected ability of a better-than-average student, $\hat{a}_{2^{n-1}-2^{n-2}}^n$ corresponds to the expected ability of a worse-than-average student, and so on (there can be a zero mass of students with a particular expected ability). S_n is not empty for $n \geq 1$ because it contains the strategy that assigns the same expected ability to all students.

Note that S_n is just the set of distributions on the set of the above $2^n - 1$ points that second-order stochastically dominate the underlying distribution of student abilities.[26] 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}). Consider now the best response correspondence $B_n(\cdot)$, which for every strategy $s \in S_n$ returns the (nonempty) set of best responses to s from the set S_n. Note that due to the continuity of payoffs, $B_n(\cdot)$ is upper hemicontinuous. Note also that for any s, the set $B_n(s)$ is convex, due to the linearity of payoffs (if strategies s_1 and s_2 are in $B_n(s)$, and thus give identical payoffs to the school, for any $\alpha \in [0, 1]$, strategy $(\alpha s_1 + (1 - \alpha)s_2)$ is also in S_n and gives the same payoff to the school, and therefore also belongs to $B_n(s)$). Thus, by Kakutani's Fixed Point Theorem there exists s_n^* such that $s_n^* \in B_n(s_n^*)$.

Take the sequence $\{s_n^*\}$ for $n \to \infty$. Since all distributions s_n^* have supports on subsets of a bounded interval $[a_L, a_H]$, this sequence has a weakly converging subsequence.

[26] Clearly, any distribution in S_n second-order stochastically dominates the underlying distribution of true student abilities. On the other hand, any distribution s that second-order stochastically dominates the underlying distribution of true student abilities can be obtained from that distribution by mixing some students together, because the underlying distribution of true student abilities is a mean-preserving spread of s.

Let s be the limit of this subsequence. Note that the payoff function of a school is continuous both in its own strategy and in the strategy of other players. Therefore, 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. 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.

The symmetry condition is sufficient for the existence of a connected equilibrium, but it is not necessary, as examples 2, 3, and 4 illustrate. However, it is not clear what a more general sufficient condition on the primitives of the model could be. As we show in the following example, a connected equilibrium may not exist, even if for any *true* ability level $a \in (a_L, a_H)$ there exists a school that has students of all *true* abilities in some ε-neighborhood of a.

Example C.1: There is mass 0.8 of students in "bad" schools, with abilities distributed uniformly on $[0, 50]$ and mass 0.2 of students in "good" schools, with abilities distributed uniformly on $[0, 150]$. There is also mass 1 of positions, distributed uniformly on $[0, 100]$.

Suppose this market has a connected equilibrium. Under truthful information revelation the desirability mapping would be concave. Hence, in the connected equilibrium, the desirability mapping would be linear, and so the observed distribution of expected abilities would be uniform on $[0, \hat{a}_H]$. The average true ability in this market is $0.8 \times 25 + 0.2 \times 75 = 35$, and the average expected ability has to be the same. Therefore, $\hat{a}_H = 70$. But no matter what mixing strategy it uses, a "good" school will produce a positive mass of students with expected ability 75 or higher—contradiction.

Hence, there is no connected equilibrium in this market. Is there another equilibrium? It turns out, there is, and moreover, the desirability mapping in it is not convex. In this equilibrium, each "bad" school reveals full information about its students, and each "good" school "compresses" the distribution of its students' abilities from the true distribution of $U[0, 150]$ to the distribution of expected ability $U[50, 100]$. The resulting desirability mapping is

$$Q(\hat{a}) = \begin{cases} \frac{8}{5}\,\hat{a}, & \text{for } \hat{a} \leq 50 \\ 80 + \frac{2}{5}\,(\hat{a} - 50), & \text{for } \hat{a} \in [50, 100], \end{cases}$$

and it is easy to check that each school behaves optimally.

The arguments in Example C.1 illustrate how to check whether a particular matching market has a connected equilibrium. The general method is as follows:

1. By Theorem 1, if a connected equilibrium exists, its desirability mapping Q has to be the same as in a market with the same aggregate distribution of abilities,

but allocated identically across schools. By Theorem C.1, that latter market has a connected equilibrium, and its desirability mapping can be constructed following the steps of the proof of Theorem 1. Construct that mapping Q.

2. Construct mapping Q_T. At every point $\hat{a}$ where Q is strictly convex, by the arguments in the proof of Theorem 1, $Q(\hat{a}) = Q_T(\hat{a})$, and in equilibrium, there is no mixing of students above and below such points.
3. Consider the (maximal) intervals on which Q is linear. Since at the boundaries of these intervals Q is strictly convex, for every such interval $[\hat{a}, \hat{b}]$, the set of students with expected abilities in that interval is the same (up to a set of measure 0) as the set of students with true abilities in that interval. Moreover, the aggregate distribution of true abilities in that interval is a mean-preserving spread of the aggregate distribution of reported abilities implied by desirability mapping Q (which is linear on that interval) and the corresponding distribution of position desirabilities. The last step is to check whether this aggregate distribution of reported abilities can be obtained as the weighted sum of mean-preserving contractions of the underlying distributions of true abilities (in range $[\hat{a}, \hat{b}]$) in schools, with weights equal to the masses of students in relevant range in those schools. If it cannot (like in Example C.1), then there cannot be a connected equilibrium. If, for every (maximal) interval, there is, then there is a connected equilibrium, because each school is indifferent over various mixings of student abilities in every interval $[\hat{a}, \hat{b}]$, since the desirability mapping there is linear, and hence every school is behaving optimally.

Finally, we would like to point out that there is a powerful force that is not captured in our model and that would tend to eliminate disconnected equilibria with nonconvex desirability mappings as in Example C.1. This force is arbitrage. In our model, we abstract away from how students get assigned to schools. In fact, as long as there exists a connected equilibrium, that does not matter. If, however, schools can compete for students and can facilitate monetary transfers between them (e.g., in the form of a high tuition and heterogeneous financial aid), then a nonconvexity in the desirability mapping would allow a school to get students from other schools, mix them together, and get a higher average payoff for them. Thus, any nonconvexity in the desirability mapping is an arbitrage opportunity, which cannot persist in equilibrium. Since, for all aggregate distributions of student abilities and position desirabilities, equilibria without this arbitrage opportunity exist for at least some allocations of students among schools (e.g., by Theorem C.1, for symmetric allocations), arbitrage will lead to one of such allocations and to an equilibrium with a convex desirability mapping.

REFERENCES

Avery, Christopher, Andrew Fairbanks, and Richard Zeckhauser. 2003. *The Early Admissions Game: Joining the Elite*. Cambridge, MA: Harvard University Press.

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): 793–902.

Becker, William E., Jr. 1982. "The Educational Process and Student Achievement Given Uncertainty in Measurement." *American Economic Review*, 72(1): 229–36.

Bergemann, Dirk, and Martin Pesendorfer. 2007. "Information Structures in Optimal Auctions." *Journal of Economic Theory*, 137(1): 580–609.

Chakraborty, Archishman, and Rick Harbaugh. 2007. "Comparative Cheap Talk." *Journal of Economic Theory*, 132(1): 70–94.

Chan, William, Hao Li, and Wing Suen. 2007. "A Signaling Theory of Grade Inflation." *International Economic Review*, 48(3): 1065–90.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite. 1992. "Social Norms, Savings Behavior, and Growth." *Journal of Political Economy*, 100(6): 1092–125.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite. 2001. "Investment and Concern for Relative Position." *Review of Economic Design*, 6(2): 241–61.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50(6): 1431–51.

Hawkins, David A., Jessica Lautz, Patricia M. McDonough, and William G. Tierney. 2005. *State of College Admission*. National Association for College Admission Counseling. Alexandria, VA, March.

Hopkins, Ed. 2010. "Job Market Signaling of Relative Position, or Becker Married to Spence." http://www.homepages.ed.ac.uk/hopkinse/signal.pdf.

Jackson, Matthew O., and Hugo F. Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." *Econometrica*, 75(1): 241–57.

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

Li, Hao, and Wing Suen. 2000. "Risk Sharing, Sorting, and Early Contracting." *Journal of Political Economy*, 108(5): 1058–91.

Lizzeri, Alessandro. 1999. "Information Revelation and Certification Intermediaries." *RAND Journal of Economics*, 30(2): 214–31.

Matthews, Steven, and Andrew Postlewaite. 1985. "Quality Testing and Disclosure." *RAND Journal of Economics*, 16(3): 328–40.

Niederle, Muriel, and Alvin E. Roth. 2003. "Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match." *Journal of Political Economy*, 111(6): 1342–52.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura. 1990. "Strategic Information Revelation." *Review of Economic Studies*, 57(1): 25–47.

Peters, Michael. 2007. "The Pre-Marital Investment Game." *Journal of Economic Theory*, 137(1): 186–213.

Peters, Michael, and Aloysius Siow. 2002. "Competing Premarital Investments." *Journal of Political Economy*, 110(3): 592–608.

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

Sattinger, Michael. 1993. "Assignment Models of the Distribution of Earnings." *Journal of Economic Literature*, 31(2): 831–80.

Suen, Wing. 2000. "A Competitive Theory of Equilibrium and Disequilibrium Unravelling in Two-Sided Matching." *RAND Journal of Economics*, 31(1): 101–20.

This article has been cited by:

1. Talia Bar, Vrinda Kadiyali, Asaf Zussman. 2012. Putting Grades in Context. *Journal of Labor Economics* **30**:2, 445-478. [CrossRef]
2. Emir Kamenica, , Matthew Gentzkow. 2011. Bayesian Persuasion. *American Economic Review* **101**:6, 2590-2615. [Abstract] [View PDF article] [PDF with links]
3. Andrew F. Daughety , Jennifer F. Reinganum. 2010. Public Goods, Social Pressure, and the Choice Between Privacy and Publicity. *American Economic Journal: Microeconomics* **2**:2, 191-221. [Abstract] [View PDF article] [PDF with links]