Notes to accompany seminar on "Applications with incomplete information"

Éva Nagypál[1]

January 23, 2003

WORK IN PROGRESS — PLEASE DO NOT CITE

Abstract

This paper studies the application behavior of agents in an environment where agents are competing for placements (in colleges, in sports teams, etc) and are allowed to submit multiple applications. It is assumed that agents have identical preferences over placements and that the agents on the other side of the market (schools, teams, etc) have identical preferences over applicant quality. With full information this is a trivial problem to study, but with incomplete information this is no longer so. In particular, two aspects of incomplete information are considered that make the application behavior in these environments an interesting decision problem to study: uncertainty about one's own quality and noise in the evaluation of applications. First I consider what happens when agents have some uncertainty about their own quality (interpreted as their ranking among the pool of applicants) and hence form beliefs about their quality and based on these they make applications. I show that with a fixed number of applications, given that the distribution characterizing the belief of the agents satisfies the monotone likelihood ratio property (MLRP), there is assortative matching on the mean belief (expected quality) of the agent in the sense that all applications are increasing in the mean belief of the agent. When I allow agents to decide how many applications to make given some application cost, this result breaks down, however. I show that, even if MLRP holds, it is possible for an agent with a better underlying quality *and* a better mean belief to secure a worse placement. I also show that the number of applications is decreasing in the dispersion of beliefs, and for high enough dispersion, agents submit no application or submit a single application for the best placement.

Next, I study what happens when a random element is added to the evaluation of the applications. I show that agents with low dispersion in their beliefs are the ones to submit multiple applications. The results thus show that both aspects of incomplete information imply that low uncertainty agents submit more applications, and hence secure better placements. In a companion paper I embed the application decision of the agents into a general equilibrium model of college applications where colleges optimally chose which students to admit, and study the properties of its equilibrium.

[1]Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, IL 60208-2600, USA, nagypal@northwestern.edu.

1 Introduction

This paper studies the application behavior of agents in an environment where agents are competing for placements (in colleges, in sports teams, etc) and are allowed to submit multiple applications. Two aspects of incomplete information are considered that make the application behavior in these environments an interesting decision problem to study: uncertainty about one's own quality and noise in the evaluation of applications. First I consider what happens when agents have some uncertainty about their own quality (interpreted as their ranking among the pool of applicants) and hence form beliefs about their quality and based on these they make applications. I show that with a fixed number of applications, given that the distribution characterizing the belief of the agents satisfies the monotone likelihood ratio property (MLRP), there is assortative matching on the mean belief (expected quality) of the agent in the sense that all applications are increasing in the mean belief of the agent. When I allow agents to decide how many applications to make given some application cost, this result breaks down, however. I show that, even if MLRP holds, it is possible for an agent with a better underlying quality *and* a better mean belief to secure a worse placement. I also show that the number of applications is decreasing in the dispersion of beliefs, and for high enough dispersion, agents submit no application or submit a single application for the best placement.

Next, I study what happens when a random element is added to the evaluation of the applications. I show that agents with low dispersion in their beliefs are the ones to submit multiple applications. This is because agents with high dispersion in their belief rationally interpret a rejection as a bad signal about their quality making subsequent applications less profitable in expectation, while agents with low dispersion rationally interpret a rejection as bad luck without much affecting the expectation about their chance of securing a placement elsewhere. The results thus show that both aspects of incomplete information imply that low uncertainty agents submit more applications, and hence secure better placements.

This paper is in the intersection of two strands of the literature. First, it is closely related to the literature on matching since it considers the decision whom to match with in a matching model with two-sided heterogeneity when meetings between partners are directed (via applications). Most matching models assume that agents are homogeneous and meetings between agents are random which means that the agents in the model cannot influence whom they will have an opportunity to match with in the subsequent period (see the search and matching literature reviewed in Pissarides (2000)). More recently, however, a literature has developed that focuses on the effects of heterogeneity on matching and search outcomes. While most papers in this literature still assume random search (for example, Burdett and Coles (1997)), there are a handful of papers that consider directed search or directed meetings with heterogeneous agents.[2] The starting point for these papers and the literature on matching with heterogeneous agents is Becker's famous result (Becker (1973)) in which he establishes that with transferable utility, given that traits of the partners are complements (i.e., given a supermodular production function), perfectly assortative matching (PAM) is the optimal matching pattern and it is achieved in equilibrium. Similarly, perfectly assortative matching is the optimal matching pattern in an environment with heterogeneous agents, directed meetings, and non-transferable utility. Perfectly assortative matching means that, in the case of the labor market, the highest quality worker matches with the highest quality firm, the next highest quality worker matches with the next highest quality firm, and so on.

This very stark prediction of perfect assortative matching is in contrast with empirical observations and intuition. There has been considerable effort in the literature to establish more general conditions under which the PAM result holds or breaks down (for example, Shimer and Smith (2000)). A recent paper emphasizing the role of coordination frictions in breaking down PAM is Shimer (2001).

The role of incomplete information has received practically no attention in this discussion. The only other extension of the matching model with heterogeneous agents, to my knowledge,

[2] The expression "directed search" is used in the literature to describe a situation in which agents can *influence* which agents they meet with and hence have the possibility to form a match with. "Directed meetings" is an extreme case of "directed search" where agents can *exactly determine* whom they have the possibility to form a match with. In a sense, there is no search nor are there search frictions in the case of "directed meetings".

that considers an incomplete information framework is Anderson and Smith (2002). Their focus is on a different aspect of incomplete information, however. In particular, they show that in a dynamic setting, PAM may fail even with a supermodular production function when matches yield not only output but also information about agents' types. They assume a much simpler form of belief heterogeneity than the one assumed in this paper (beliefs are binary in their model) and do not consider multiple applications, hence the questions asked in this paper about the effects of incomplete information cannot be studied within their framework.

An important issue in matching models is whether utility between the parties in the match is transferable, for example via wages, or not. In the model studied I assume that utility is not transferable between the partners in a match[3], hence the model is more appropriate to study markets where this assumption is more plausible, such as the market for college admissions. Thus the model is phrased is terms of college application decisions, even though its implications are more broadly applicable to matching markets with multiple applications. While the assumption of non-transferable utility does not completely characterize the college admission process, since colleges actively price discriminate by using financial aid policies, most admission policies still claim to be "need-blind". The extent of price discrimination is also limited by the fact that colleges have set tuition rates, and at their set tuition rates the demand for selective colleges still far exceeds supply, implying that price discrimination and hence transferability of utility is far from perfect between colleges and their students.

This brings us to the second strand of the literature to which this paper is closely related, that on college applications. The decision to apply to college is the first major economic decision most young adults undertake as they embark upon their adult life. The stakes are perceived to be large, as is manifested by the tremendous attention paid to the admission policies of the most selective colleges by students, their parents, and commentators at large. The evidence available on labor market performance also points in the direction of substantial labor market payoffs to attending a selective college (Brewer, Eide, and Ehrenberg (1999)).

[3] A good comparison of models with non-transferable utility to those with transferable utility can be found in Burdett and Coles (1999).

Despite the importance that this decision has in the early stage of an individual's economic life, the theoretical literature studying the college application decision is not very large. The most prominent strand of this literature is the game theory and mechanism design literature following Gale and Shapley's seminal work on the so-called "college admission problem" (see Gale and Shapley (1962)). In the college admission problem students and colleges have diverse preferences over which colleges to attend and which students to admit and the task is to design a mechanism that results in an individually rational and stable outcome (i.e., an outcome in which no agent would prefer to be unmatched and in which no two agents who are not matched to each other would both prefer to be matched with each other).

While the heterogeneity of preferences in a complete information environment is certainly an important aspect of the college application decision, issues relating to incomplete information are equally important. There are two important sources of incomplete information in the college application process. First, high school graduates generally do not fully know how well they rank in the general pool of college applicants. Of course, they might know their high school ranking and if they take the SAT (which is itself already a cost of applying to college) they know their SAT scores and the corresponding percentiles, but these pieces of information do not fully reveal the quality of a college applicant. Hence, students face some uncertainty regarding their own quality. Moreover, even if students were fully aware of their own ranking within the pool of college applicants and had the same preferences over colleges, one would still expect them not to sort perfectly on quality to the different quality colleges, due to the presence of "noise" in the admission process. In other words, while the assessment of a student's ranking within the college applicant pool across colleges is certainly correlated, this correlation is far from perfect.

I consider a model where high school graduates (henceforth students) are heterogeneous in their belief about their own quality. This quality can be thought of as the ranking of the student among the pool of college applicants. Students' beliefs are characterized by a distribution function parameterized by its mean and variance. These students then decide which colleges, if any, to apply to. Applying to college has a fixed cost and each application has a variable cost. Each application gains the possibility for the student of being admitted by the college she applied to. Colleges are of a finite number of types based on their reputation, and there are many colleges of each type. Students have the same preferences over colleges, i.e., they all prefer colleges with better reputations, and this is the only attribute of colleges that they care about. Better reputation colleges have higher admission standards, meaning that they have a higher quality cutoff for accepting students. I study two versions of the model. In the first, colleges observe the actual quality (absolute ranking) of an applicant, while in the second they do not observe the actual quality (absolute ranking) of an applicant, instead they observe a noisy realization of it.

In a companion paper I embed the application decision of the agents into a general equilibrium model of college applications where colleges optimally chose which students to admit, and study the properties of its equilibrium.

The paper proceeds as follows.

2 A simple model with uncertainty about own quality

There are two colleges in the economy. The two colleges differ in terms of the payoff a student gets from attending them, i.e., their desirability (which I will call *reputation* from now on). The high-reputation college has reputation η_h, while the low-reputation college has reputation $\eta_l < \eta_h$. The threshold quality requirement of the high-reputation college is q_h, while the threshold quality requirement of the low-reputation college is $q_l < q_h$.

Each student in the economy has a distinct quality, μ, that is drawn from a normal distribution with mean 0 and variance σ_μ^2. The student does not know her own quality. Instead, she gets a signal about her quality, $\hat{\mu} = \mu + \varepsilon$, where ε has a distribution $N(0, \sigma_\varepsilon^2)$. Once the student receives this signal, she updates her belief about her quality, and her posterior belief will be normal with mean

$$\tilde{\mu} = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varepsilon^2}(\mu + \varepsilon),$$

and variance

$$\tilde{\sigma}^2 = \frac{\sigma_\varepsilon^2 \sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varepsilon^2}.$$

Since there is a one-to-one correspondence between σ_ε^2 and $\tilde{\sigma}^2$, I can treat the latter as the primitive of the model, which will turn out to be convenient.

Once the student updates her belief about her quality, she decides whether to apply to any of the two colleges in the economy. When making this decision, the student is maximizing her expected payoff. There is a cost of applying to college, $c \geq 0$ per application. Since there are two colleges in the economy, the student submits at most two applications. (There is no benefit from submitting multiple applications to the same college.) If the student's quality is above the threshold quality requirement of a college she applied to, then she is accepted by that college. Finally, the student decides which college to enroll of the ones that accepted her, and if the student enrolls college j, then she gets a payoff of $\eta_j > 0$. If the student does not enroll college, her payoff is 0.

Consider the application policy of the students given the threshold $q_l < q_h$ of the two colleges. (These threshold qualities are considered exogenous in this paper. For an equilibrium derivation of the threshold qualities, see the companion paper.) A student with posterior belief $(\tilde{\mu}, \tilde{\sigma})$ has four options: she can choose not to apply to any college, she can apply to the low-reputation college, to the high-reputation college, or to both colleges. The optimal policy of a student with any belief can be described by comparing the above alternatives, two at a time, and establishing the values of $(\tilde{\mu}, \tilde{\sigma})$ for which each is preferred.

2.1 Pairwise comparison of alternative policies

2.1.1 Applying to a single college versus not applying

If the student with belief $(\tilde{\mu}, \tilde{\sigma})$ applies to the low-reputation college, then her expected payoff is

$$\bar{F}\left(\frac{q_l - \tilde{\mu}}{\tilde{\sigma}}\right)\eta_l - c, \tag{1}$$

where $\bar{F}$ is the survival function of the standard normal distribution, i.e., $\bar{F} = 1 - F$, where F is the cumulative density function of the standard normal distribution. One necessary condition for this expected payoff to be higher than 0, the payoff from not applying, is that $c < \eta_l$. This assumption is maintained throughout, since it simply means that the cost of applying to the low-reputation college is outweighed by the payoff from attending that college. Then, given this assumption, the expected payoff is higher than 0 if and only if

$$\tilde{\mu} \geq q_l - F^{-1}\left(1 - \frac{c}{\eta_l}\right)\tilde{\sigma}.$$

This relationship defines a curve in $(\tilde{\mu}, \tilde{\sigma})$ space that separates the region where it is better to apply to the low-reputation college from the one where it is better not to apply. This curve is either downward- or upward-sloping, depending on the sign of $F^{-1}\left(1 - \frac{c}{\eta_l}\right)$. This curve is denoted m_{nl} in Figure 2.

Similarly, if the student with belief $(\tilde{\mu}, \tilde{\sigma})$ applies to the high-reputation college, then her expected payoff is

$$\bar{F}\left(\frac{q_h - \tilde{\mu}}{\tilde{\sigma}}\right)\eta_h - c. \tag{2}$$

Given that $c < \eta_l < \eta_h$, this expected payoff is higher than 0 if and only if

$$\tilde{\mu} \geq q_h - F^{-1}\left(1 - \frac{c}{\eta_h}\right)\tilde{\sigma}.$$

Similarly to the m_{nl} curve, this relationship defines the m_{nh} curve in $(\tilde{\mu}, \tilde{\sigma})$ space, as demonstrated in Figure 2. The slope of the m_{nh} curve is lower than the slope of the m_{nl} curve, since $\eta_h > \eta_l$.

2.1.2 Applying to the low-reputation college versus applying to both colleges

If the student with belief $(\tilde{\mu}, \tilde{\sigma})$ applies to both colleges, then her expected payoff is

$$\bar{F}\left(\frac{q_h - \tilde{\mu}}{\tilde{\sigma}}\right)\eta_h + \left[\bar{F}\left(\frac{q_l - \tilde{\mu}}{\tilde{\sigma}}\right) - \bar{F}\left(\frac{q_h - \tilde{\mu}}{\tilde{\sigma}}\right)\right]\eta_l - 2c. \tag{3}$$

Comparing this with the payoff in Equation (1) from applying to just the low-reputation college, we can see that it is worth applying to both colleges instead of applying to just the low-reputation college if and only if

$$\bar{F}\left(\frac{q_h - \tilde{\mu}}{\tilde{\sigma}}\right)(\eta_h - \eta_l) - c \geq 0.$$

This holds if $c < \eta_h - \eta_l$ and

$$\tilde{\mu} \geq q_h - F^{-1}\left(1 - \frac{c}{\eta_h - \eta_l}\right)\tilde{\sigma}.$$

This relationship defines the m_{bl} curve in $(\tilde{\mu}, \tilde{\sigma})$ space. Its slope is clearly higher than that of m_{nh} (as in Figure 2).

2.1.3 Applying to the high-reputation college versus applying to both colleges

Comparing the payoff from applying to both colleges in Equation (3) to that from applying to the high-reputation college in Equation (2), we can see that it is worth applying to both colleges instead of applying to just the high-reputation college if

$$\bar{F}\left(\frac{q_l - \tilde{\mu}}{\tilde{\sigma}}\right) - \bar{F}\left(\frac{q_h - \tilde{\mu}}{\tilde{\sigma}}\right) \geq \frac{c}{\eta_l}.$$

This relationship defines the m_{bh} curve in $(\tilde{\mu}, \tilde{\sigma})$ space in Figure 2.

2.1.4 Applying to the high-reputation versus the low-reputation college

Comparing the payoff from applying to the high-reputation college in Equation (2) versus that from applying to the low-reputation college in Equation (1), we can see that it is worth applying to the high-reputation college instead of applying to the low-reputation college if and only if

$$\frac{\bar{F}\left(\frac{q_h-\tilde{\mu}}{\tilde{\sigma}}\right)}{\bar{F}\left(\frac{q_l-\tilde{\mu}}{\tilde{\sigma}}\right)} \geq \frac{\eta_l}{\eta_h}.$$

This relationship defines the m_{lh} curve in $(\tilde{\mu}, \tilde{\sigma})$ space in Figure 2.

2.2 The role of the application cost

Based on the value of the application cost, c, we can distinguish two cases based on the value of the model's exogenous parameters (notably, not based on the value of q_l and q_h.) First, if $(\eta_h - \eta_l)\frac{\eta_l}{\eta_h} \leq c$, then the m_{bl} and m_{bh} curves do not intersect (since that would imply $\bar{F}\left(\frac{q_l-\tilde{\mu}}{\tilde{\sigma}}\right) \geq 1$, which is not possible for $\tilde{\sigma} > 0$.) This case is represented in Figure 1. Note that, when $\eta_h - \eta_l \leq c$, then there is no part of the $(\tilde{\mu}, \tilde{\sigma})$ space in which applying to both colleges is preferred to applying to just the low-reputation college (i.e., there is no m_{bl} curve), but this is qualitatively the same as when there is an m_{bl} curve that does not intersect the m_{bh} curve. The empty circles represent $(\tilde{\mu}, \tilde{\sigma})$ combinations for which it is preferable to apply to the low-reputation college, while the small dots represent posterior beliefs for which it is preferable to apply to the high-reputation college. (Circles with dots in them represent posterior beliefs for which it is preferable to apply to both colleges, though there are no such beliefs in this case.)

This is the case when the cost of the second application is prohibitively expensive so that a second application is never submitted. While this case features no multiple applications, it demonstrates several interesting results. First, for a given posterior standard deviation, the PAM result holds in the sense that there is PAM along the posterior means of the agents, those with high posterior means (good signals) match with the high-reputation college, while those with low posterior means (bad signals) match with the low-reputation college. (An equivalent result holds in the static version of the model of Anderson and Smith (2002).)

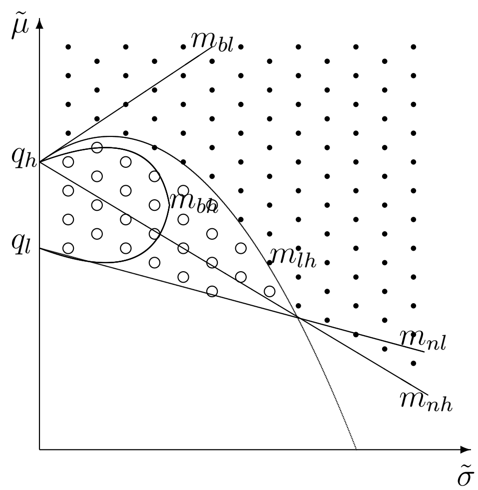

Figure 1: Optimal application policy of students as a function of posterior mean and standard deviation in the case when $(\eta_h - \eta_l)\frac{\eta_l}{\eta_h} \leq c_2$.

More interestingly, however, we can see from Figure 1 that students with the same posterior mean but different posterior standard deviation might follow different application policies. For very high values of the posterior standard deviation, students will "take long shots" and apply to the high-reputation college or not apply, without considering to apply to the low-reputation college for any value of the posterior mean. For lower values of the posterior standard deviation, students with intermediate values of the posterior mean will "play it safe" and apply to the low-reputation college. This means that some students with the same posterior mean will follow different application policies and potentially will enroll different colleges. Thus Becker's result no longer holds in the simple sense of PAM on posterior means.

The second case, when $c < (\eta_h - \eta_l)\frac{\eta_l}{\eta_h}$, is displayed in Figure 2. In this case the m_{bl} and m_{bh} curves intersect. Once again, the empty circles represent $(\tilde{\mu}, \tilde{\sigma})$ combinations for which it is preferable to apply to the low-reputation college, the small dots represent posterior beliefs for which it is preferable to apply to the high-reputation college and the circles with dots in them represent posterior beliefs for which it is preferable to apply to both colleges.

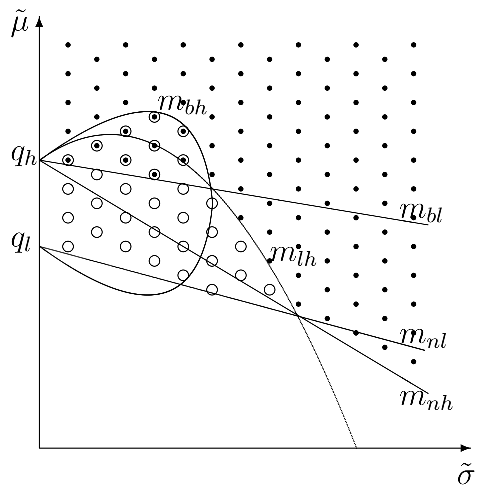

Figure 2: Optimal application policy of students as a function of posterior mean and standard deviation in the case when $c < (\eta_h - \eta_l)\frac{\eta_l}{\eta_h}$.

This case is more interesting than the previous one in the sense that it features multiple applications for some posterior beliefs. The $(\tilde{\mu}, \tilde{\sigma})$ combinations for which students will choose to apply to both colleges are relatively low values of the posterior standard deviation coupled with posterior means around the cutoff of the high-reputation college. These students have a quite precise idea about their own quality and realize that being so close to the cutoff of the high-reputation college, they run the risk of "just not making it", i.e., ending up with a quality that is just below the cutoff.

3 The general model

There are a continuum of colleges in the economy that differ in their threshold quality requirement, q. Moreover, the colleges differ in terms of the payoff a student gets from attending them, their desirability, denoted by $\eta(q)$. It is assumed that, over a relevant range, $\eta(q)$ is increasing, meaning that colleges that are more stringent in their quality requirement provide a higher payoff to students attending them.[4]

Assumption 1. *$\eta(q)$ is continuous with $\eta(q) = \underline{\eta} = 0$ on $q \in (-\infty, \underline{q})$, $\eta'(q) > 0$ on $q \in (\underline{q}, \bar{q})$, and $\eta(q) = \bar{\eta} > 0$ on $q \in (\bar{q}, \infty)$ for some $\underline{q}$ and $\bar{q}$.*

Each student has a belief about her own quality characterized by the distribution function $F(\mu)$. Instead of specifying how a student arrives to having the beliefs she has, I make more general assumptions on the distribution function $F(\mu)$ that allow me to derive results about the application behavior of students. $F(\mu)$ can, of course, be the result of Bayesian updating as in the above examples, but this need not be the case.

Assumption 2. *$F(\mu)$ is a probability distribution function with support on $(-\infty, \infty)$. $F(\mu)$ is continuous, strictly increasing and continuously differentiable.*

Assumption 3. *$F(\mu)$ belongs to a class of distribution functions parameterized by $\tilde{\mu}$ and $\tilde{\sigma}$ such that*

$$F(\mu \mid \tilde{\mu}, \tilde{\sigma}) = F\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}} \mid 0, 1\right).$$

Assumption 4. *$F(\mu \mid \tilde{\mu}, \tilde{\sigma})$ satisfies the strict monotone likelihood ratio property, i.e., for any $\mu_1 < \mu_2$, $\frac{\bar{F}\left(\frac{\mu_2 - \tilde{\mu}}{\tilde{\sigma}}\right)}{\bar{F}\left(\frac{\mu_1 - \tilde{\mu}}{\tilde{\sigma}}\right)}$ is strictly increasing in $\tilde{\mu}$.*

Next, I consider two alternative assumptions about the application behavior of student. In Section 3.1, I consider the case when students can submit a fixed number of applications,

[4] In a companion paper, I consider the optimizing behavior of colleges when setting their admission standards and the resulting equilibrium. I show there that, taking the payoffs of the colleges — the distribution of η — as the primitives of the model, the equilibrium admission standard, $q(\eta)$, is in fact increasing in η. The inverse notation used in this paper is more convenient for the problem at hand.

while in 3.2 I endogenize the application intensity by allowing students to submit any number of applications given some application cost.

3.1 Fixed number of applications

Assumption 5A. *A student can submit a fixed N number of applications.*

Given that $\eta'(q) > 0$ on $(\underline{q}, \bar{q})$, it immediately follows that all student will submit the maximum allowed number of applications. Let us introduce the following notation. Let us denote the ordered optimal acceptance set of a student with belief $F(\tilde{\mu}, \tilde{\sigma})$ by $A(\tilde{\mu}, \tilde{\sigma})$ and its i^{th} element by $q_i^*(\tilde{\mu}, \tilde{\sigma})$. Moreover, let us denote the highest element of $A(\tilde{\mu}, \tilde{\sigma})$ by $h(\tilde{\mu}, \tilde{\sigma})$ and its lowest by $l(\tilde{\mu}, \tilde{\sigma})$.

Theorem 1 (Assortative matching given fixed number of applications). *Given Assumptions 1 through 4 and 5A, for a given $\tilde{\sigma}$, $q_i^*(\tilde{\mu}, \tilde{\sigma})$ (and hence $h(\tilde{\mu}, \tilde{\sigma})$ and $l(\tilde{\mu}, \tilde{\sigma})$) is increasing in $\tilde{\mu}$.*

Proof of Theorem 1. Let us fix $\tilde{\sigma} = 1$ (it is easy to show that this simply makes the notation a little less cumbersome and is without loss of generality). Given a fixed N number of applications, the objective of the student is to choose $q_1 < q_2 < ... < q_N$ to maximize

$$\sum_{i=1}^{N} \overline{F}\left(q_i - \tilde{\mu}\right)\left[\eta(q_i) - \eta(q_{i-1})\right].$$

(For notational convenience let us use the notation $q_0 = -\infty$ and $q_{N+1} = \infty$, so that $\eta(q_0) = 0$ and $\overline{F}\left(q_{N+1} - \tilde{\mu}\right) = 0$.) The optimal choices $\mathbf{q}^* = [q_1^*, ..., q_N^*]$ (the arguments of which are suppressed) satisfy the necessary first-order conditions

$$\left[\overline{F}\left(q_i - \tilde{\mu}\right) - \overline{F}\left(q_{i+1} - \tilde{\mu}\right)\right]\eta'(q_i) - f\left(q_i - \tilde{\mu}\right)\left[\eta(q_i) - \eta(q_{i-1})\right] = 0.$$

The second-order condition for a maximum requires that the Hessian evaluated at $\mathbf{q}^*$ be negative definite:

$$H = \begin{bmatrix} d_1 & o_1 & 0 & 0 & . & . & . \\ o_1 & d_2 & o_2 & 0 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & 0 & o_{i-1} & d_i & o_i & 0 & . \\ . & . & . & . & . & . & . \\ . & . & 0 & 0 & o_{n-2} & d_{n-1} & o_{n-1} \\ . & . & . & 0 & 0 & o_{n-1} & d_n \end{bmatrix},$$

where

$$d_i = \left[\overline{F}\left(q_i^* - \tilde{\mu}\right) - \overline{F}\left(q_{i+1}^* - \tilde{\mu}\right)\right]\eta''(q_i^*) - f'\left(q_i^* - \tilde{\mu}\right)\left[\eta(q_i^*) - \eta(q_{i-1}^*)\right] - 2f\left(q_i^* - \tilde{\mu}\right)\eta'(q_i^*),$$

and

$$o_i = f(q_{i+1}^* - \tilde{\mu})\eta'(q_i^*).$$

Notice that o_i needs to be strictly positive since $f(q_i^*) > 0$ by Assumption 2 and $\eta'(q_i^*) > 0$, $i = 1, ..., n-1$ since, given Assumption 1, the optimal school choices feature $\underline{q} \leq q_1^* < ... < q_{N-1}^* < q_N^* \leq \bar{q}$.

Totally differentiating the N first-order conditions with respect to $\tilde{\mu}$ and suppressing $\tilde{\mu}$ in what follows, we get

$$[f(q_i^*) - f(q_{i+1}^*)]\eta'(q_i^*) + f'(q_i^*)[\eta(q_i^*) - \eta(q_{i-1}^*)] + o_{i-1}\frac{dq_{i-1}^*}{d\tilde{\mu}} + d_i\frac{dq_i^*}{d\tilde{\mu}} + o_i\frac{dq_{i+1}^*}{d\tilde{\mu}} = 0,$$

or in matrix notation

$$p + Hg = 0,$$

where p is a vector of length N with general element

$$p_i = \left[f(q_i^*) - f(q_{i+1}^*)\right] \eta'(q_i^*) + f'(q_i^*)[\eta(q_i^*) - \eta(q_{i-1}^*)]$$

and g is a vector of length N with general element

$$g_i = \frac{dq_i^*}{d\tilde{\mu}}.$$

Multiplying by the inverse of H and rearranging, we get an expression for the gradient vector

$$g = -H^{-1}p. \tag{4}$$

First notice that every element of p is positive given a distribution function satisfying strict MLRP, because

$$\begin{aligned}
\frac{p_i}{f(q_i)[\eta(q_i) - \eta(q_{i-1})]} &= \frac{[f(q_i) - f(q_{i+1})]\eta'(q_i)}{f(q_i)[\eta(q_i) - \eta(q_{i-1})]} + \frac{f'(q_i)}{f(q_i)} = \frac{f(q_i) - f(q_{i+1})}{\overline{F}(q_i) - \overline{F}(q_{i+1})} + \frac{f'(q_i)}{f(q_i)} = \\
&= \frac{-\int_{q_i}^{q_{i+1}} f'(q)dq}{\int_{q_i}^{q_{i+1}} f(q)dq} + \frac{f'(q_i)}{f(q_i)} = \frac{-\int_{q_i}^{q_{i+1}} \frac{f'(q)}{f(q)} f(q)dq}{\int_{q_i}^{q_{i+1}} f(q)dq} + \frac{f'(q_i)}{f(q_i)} > \\
&> \frac{-\frac{f'(q_i)}{f(q_i)} \int_{q_i}^{q_{i+1}} f(q)dq}{\int_{q_i}^{q_{i+1}} f(q)dq} + \frac{f'(q_i)}{f(q_i)} = 0,
\end{aligned}$$

where the second equality follows from the first-order condition, while the inequality follows from the fact that a distribution function satisfies strict MLRP iff $\frac{f'(x)}{f(x)}$ is strictly declining.

To show that all the elements of g are positive, it is then sufficient to show that all the elements of H^{-1} are negative. Since H is negative definite, we know that it is true for an arbitrary vector $k = [k_1, ..., k_N]'$ that

$$k'Hk < 0, \qquad \text{or} \qquad \sum_{i=1}^{N} d_i k_i^2 + \sum_{i=1}^{N-1} 2o_i k_i k_{i+1} < 0.$$

Now consider the j^{th} column of H^{-1}, $[h_{1j}, ..., h_{Nj}]'$. From the fact that H^{-1} is the inverse of

H, we know that

$$\begin{aligned}
d_1 h_{1j} + o_1 h_{2j} &= 0 \\
o_{N-1} h_{N-1j} + d_N h_{Nj} &= 0. \\
o_{i-1} h_{i-1j} + d_i h_{ij} + o_j h_{i+1j} &= 0 \qquad \text{for} \quad i = 2, ..., j-1, j+1, .., N-1, \text{ and} \\
o_{j-1} h_{j-1j} + d_j h_{jj} + o_j h_{j+1j} &= 1 \qquad (5)
\end{aligned}$$

Now consider the vector $[h_{1j}, ..., h_{kj}, 0, .., 0]'$ for any $k < j$. Since H is negative definite, we know that

$$\sum_{i=1}^{k} d_i h_{ij}^2 + \sum_{i=1}^{k-1} 2 o_i h_{ij} h_{i+1j} < 0,$$

or

$$\begin{aligned}
& h_{1j}(d_1 h_{1j} + o_1 h_{2j}) + h_{2j}(o_1 h_{1j} + d_2 h_{2j} + o_2 h_{3j}) + ...+ \\
+ \quad & h_{k-1j}(o_{k-2} h_{k-2j} + d_{k-1} h_{k-1j} + o_{k-1} h_{kj}) + h_{kj}(o_{k-1} h_{k-1j} + d_k h_{kj}) < 0,
\end{aligned}$$

which together with (5) implies that

$$-h_{kj} o_k h_{k+1j} < 0.$$

Since o_k is positive, this means that h_{kj} and h_{k+1j} have the same sign. Since this holds for any $k < j$, this means that h_{1j} through h_{jj} all have the same sign.

Now considering the vector $[0, ..., 0, h_{kj}, .., h_{Nj}]'$ for any $k > j$, we can similarly establish that h_{jj} through h_{Nj} all have the same sign, hence all the elements of column j have the same sign. Since H^{-1} is symmetric (it is the inverse of a symmetric matrix), this means that all the elements of H^{-1} have the same sign. Finally, since H^{-1} is negative definite (it is the inverse of a negative definite matrix), this means that all the elements of H^{-1} are negative. Given (4), this in turn implies that all the elements of g are positive. (If the assumption of strict MLRP were to be substituted with the assumption of MLRP, then the same argument would show that the elements of g are non-negative, since then the elements of p need not be strictly positive, just non-negative.) $\square$

3.2 Endogenous number of applications

Now let us assume that instead of submitting a fixed number of applications, a student can submit any number of applications given some application cost.

Assumption 5B. *A student can submit any number of applications at a cost of c per application.*

Theorem 2 (Non-assortative matching given an endogenous number of applications). *Given Assumptions 1 through 4 and 5B, for a given $\tilde{\sigma}$, neither $h(\tilde{\mu}, \tilde{\sigma})$ nor $l(\tilde{\mu}, \tilde{\sigma})$ need to be increasing in $\tilde{\mu}$.*

This means that even if the assumption of monotone likelihood ratio property holds, which was the key to assortative matching with one or a fixed number of applications, once the application intensity is internalized, assortative matching need not hold, not even in the weak sense of the highest or lowest quality application increasing in the belief of the student.

Theorem 2 implies that — in a Bayesian updating environment — it is possible that student A is of better underlying quality than student B *and* student A gets a better signal about quality than student B, nonetheless student B ends up enrolling a better quality college than student A. To show that this is the case, let us consider the following example.

Example 1. *Let the application cost be equal to $c = .117$, let the distribution function be normal with mean $\tilde{m}u$ and variance 1 and let the payoff function be piece-wise linear.*

$$\eta(q) = \begin{cases} 0 & q \in (-\infty, -4/3] \\ 1 + .75q & q \in (-4/3, 0] \\ 1 + .35q & q \in (0, \infty) \end{cases}$$

Figure 3 plots difference between the value of optimally applying to college given that the student submits one application and the value applying to college given that the student *submits two applications. (In this example, it is never beneficial for the student to submit more than two applications.)*

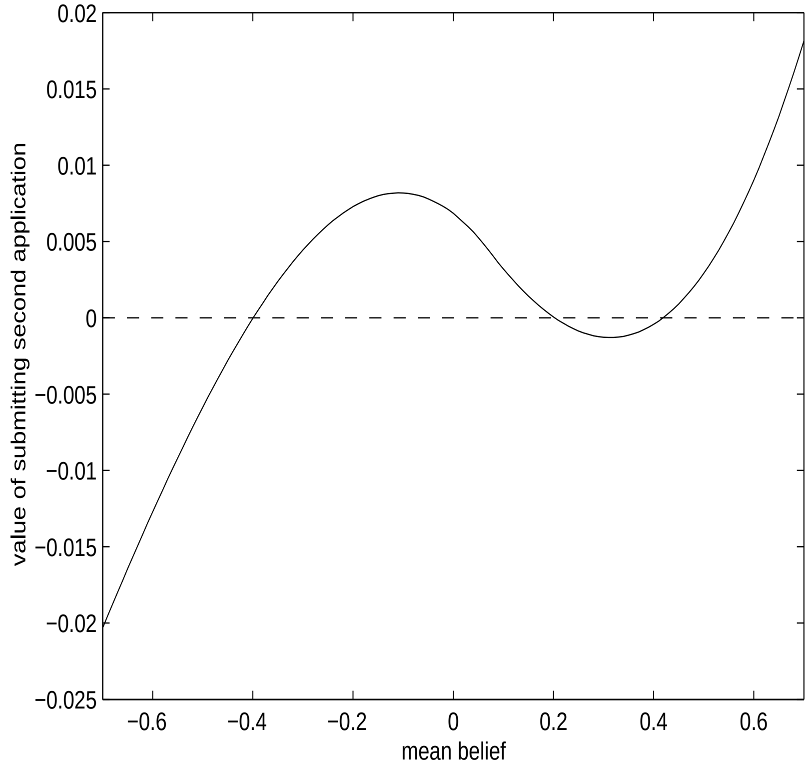

Figure 3: Value of submitting a second application as a function of the mean of the student's belief.

Figure 4 plots the optimal application behavior of a student as a function of $\tilde{\mu}$.

Now let us assume that the underlying quality distribution in the population is $N(0, 2)$*, and the student receive signals that are contaminated by an error with a distribution* $N(0, 2)$*. Now let us assume that student A has inherent quality* $\mu_A = .30$*, while student B has inherent quality* $\mu_B = .20 < \mu_A$*. Student A receives a signal* $\mu_A + \epsilon_A = .30 + .30 = .60$ *and — using the simple Bayesian updating formulas above — updates her belief to* $N(.30, 1)$*, while student B receives a signal* $\mu_B + \epsilon_A = .20 + .20 = .40$ *and updates his belief to* $N(.20, 1)$*. This implies that student A will optimally apply and get into school* $q_A = 0$*, while student B will optimally apply to schools* $q_{B1} = -.51$ *and* $q_{B2} = .174$*, get into both and enroll school* $q_{B2} = .174$*, a better school than the one attended by student A.*

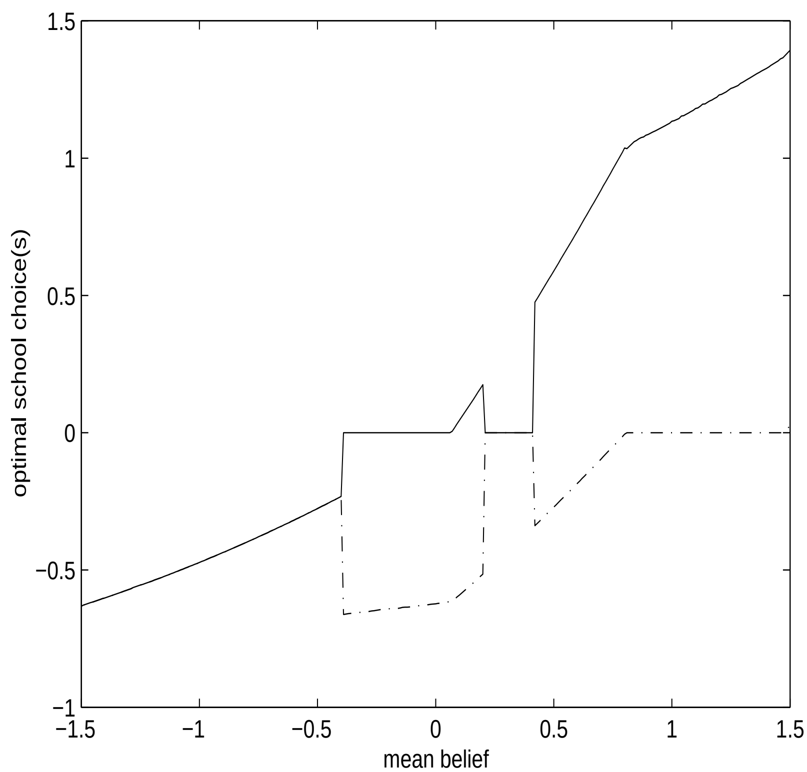

Figure 4: Optimal choice of school(s) as a function of the mean of the student's belief.

Theorem 3. *There exists $\hat{\sigma} < \infty$ such that for all $\tilde{\mu} \in \boldsymbol{R}$ and $\tilde{\sigma} > \hat{\sigma}$ the acceptance set is $A(\tilde{\mu}, \tilde{\sigma}) = \{\bar{q}\}$ or $A(\tilde{\mu}, \tilde{\sigma}) = \emptyset$.*

4 Noise in the evaluation of applications

Now consider the following alternative scenario. There are again two colleges in the economy. These two colleges now, however, are exactly the same in terms of the payoff a student gets from attending them. The threshold quality requirement of each college is q. The information structure on the side of the student is the same as in Section 2, hence the student's belief about her quality is normal with mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$.

Once the student updates her belief about her quality, she decides how many of the colleges to apply to, if any. If a student applies to a college, then that college receives a signal about the student $\hat{\mu} = \mu + \nu_j$, where ν_j is independent across colleges and students and is distributed $N(0, 1)$.[5] If this signal is above the threshold q, then the student is accepted, otherwise she is rejected.

Once again, the student is maximizing her expected payoff and again the cost per application is c. Since there are two colleges in the economy, the student submits at most two applications. A student with posterior belief $(\tilde{\mu}, \tilde{\sigma})$ has three options: she can choose not to apply to any college, she can apply one college or to two colleges. The optimal policy of a student with any belief can be described by comparing the above alternatives, two at a time, and establishing the values of $(\tilde{\mu}, \tilde{\sigma})$ for which each is preferred.

Intuitively, the decision of students to make multiple applications is driven by the fact that the extra payoff from applying to an additional college of the same quality is higher if the correlation between the signals that the colleges receive is lower. In the extreme, if the signals are perfectly correlated, then there is no reason for students to submit multiple applications.

The unconditional correlation between the signals that the two colleges receive is $\frac{\sigma_\mu^2}{\sigma_\mu^2 + 1}$.

[5]Assuming that the noise has a variance of 1 is without loss of generality, since what is important in the inference process is its relative magnitude compared to $\sigma\mu$, which is allowed to take on any value.

The correlation conditional on the information available to the student is lower, however, it is

$$\frac{\tilde{\sigma}_\mu^2}{\tilde{\sigma}_\mu^2 + 1},$$

which is clearly increasing in $\tilde{\sigma}_\mu^2$. Hence students with higher precision about their own quality (lower posterior variance) will perceive the correlation between the signals about them that the colleges receive to be lower. This means that it is these students that will benefit more from making multiple applications. To put it differently, if a college rejects a high uncertainty student with a given posterior mean, she rationally interprets it as a signal that she is in fact not good enough to get into that quality college, while if a college rejects a low uncertainty student with the same posterior mean, she rationally interprets it as "bad luck", i.e. a low realization of the noise term ν_j. This is because the admission decision of a college carries lower information content for a well-informed student.

The probability of getting into college if making just one applications is $P_1(\tilde{\mu}, \tilde{\sigma})$, where this is equal to

$$P_1(\mu+\nu_1 > q \mid \tilde{\mu}, \tilde{\sigma}) = \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right) P(\nu_1 > q-\mu) d\mu = \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right) \bar{F}\left(q - \mu\right) d\mu.$$

Proposition 4. *$P_1(\tilde{\mu}, \tilde{\sigma})$ is increasing in $\tilde{\mu}$ and is increasing in $\tilde{\sigma}$ if and only if $\tilde{\mu} \leq q$.*

Proof: See Appendix.

Proposition 4 implies that in $(\tilde{\sigma}, \tilde{\mu})$-space, the indifference curve between applying to a single college versus not applying at all is monotone. If it intersects the $\tilde{\mu}$ axis below q, which occurs if $P_1(q, 0)\eta - c \geq 0$, i.e., if $c \leq \frac{\eta}{2}$, then the indifference curve is downward-sloping. On the contrary, if it intersects the $\tilde{\mu}$ axis above q, which occurs if $c \geq \frac{\eta}{2}$, then the indifference curve is upward-sloping.

The probability of getting into college if making two applications is $P_2(\tilde{\mu}, \tilde{\sigma})$, where this is equal to

$$P_2(\mu+\nu_1 > q \text{ or } \mu+\nu_2 > q \mid \tilde{\mu},\tilde{\sigma}) =$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-P(\nu_1 < q-\mu)P(\nu_2 < q-\mu)\right] d\mu =$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-F\left(q-\mu\right)^2\right] d\mu.$$

It is preferable to apply to two colleges as opposed to just one if the payoff from doing so is higher, i.e., when

$$P_2(\tilde{\mu},\tilde{\sigma})\eta - 2c \geq P_1(\tilde{\mu},\tilde{\sigma})\eta - c,$$

or

$$P_2(\tilde{\mu},\tilde{\sigma}) - P_1(\tilde{\mu},\tilde{\sigma}) = \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right) F\left(q-\mu\right)\bar{F}\left(q-\mu\right) d\mu \geq \frac{c}{\eta}.$$

Proposition 5. *$P_2(\tilde{\mu},\tilde{\sigma}) - P_1(\tilde{\mu},\tilde{\sigma})$ is increasing in $\tilde{\mu}$ if and only if $\tilde{\mu} \leq q$. Moreover, for each $\tilde{\mu}$, there exists a $0 \leq \hat{\sigma}(\tilde{\mu}) \leq 2 \mid q-\tilde{\mu} \mid$ such that $P_2(\tilde{\mu},\tilde{\sigma}) - P_1(\tilde{\mu},\tilde{\sigma})$ is decreasing in $\tilde{\sigma}$ for any $\tilde{\sigma} \geq \hat{\sigma}(\tilde{\mu})$.*

Proof: See Appendix.

Proposition 5 implies that if any student applies to two colleges, it will be the ones with posterior means close to the cutoff of the colleges and with low posterior variances. Moreover, if for some student finds it profitable to apply to two colleges, then the student who is certain about her quality and has a quality equal to the cutoff quality of the colleges must find it profitable to apply to two colleges.[6] This student finds two applications profitable if $(P_2(q,0) - P_1(q,0))\,\eta \geq c$, or $c \leq \frac{\eta}{4}$.

Hence, based on the magnitude of the application costs, we can distinguish between three cases. When $c \geq \frac{\eta}{2}$, the indifference curve between applying to one college and not applying

[6]This is because from Proposition 5 it follows that $P_2(\tilde{\mu},\tilde{\sigma}) - P_1(\tilde{\mu},\tilde{\sigma}) \leq P_2(q,\tilde{\sigma}) - P_1(q,\tilde{\sigma}) \leq P_2(q,0) - P_1(q,0)$.

is upward-sloping and the application cost is so high that no student finds it preferable to apply to two colleges as opposed to applying to just one. When $\frac{\eta}{4} \leq c < \frac{\eta}{2}$ the same indifference curve is downward-sloping though there are still no multiple applications. This latter case is depicted in Figure 5.

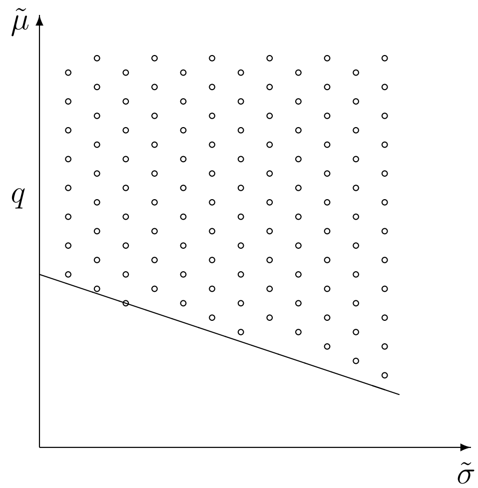

Figure 5: Optimal application policy of students as a function of posterior mean and standard deviation when there is noise in the decision process in the case when $\frac{\eta}{4} \leq c < \frac{\eta}{2}$.

In the last case, some students find it preferable to apply to two colleges as opposed to applying to just one, i.e., $c < \frac{\eta}{4}$. This is the case depicted in Figure 6.

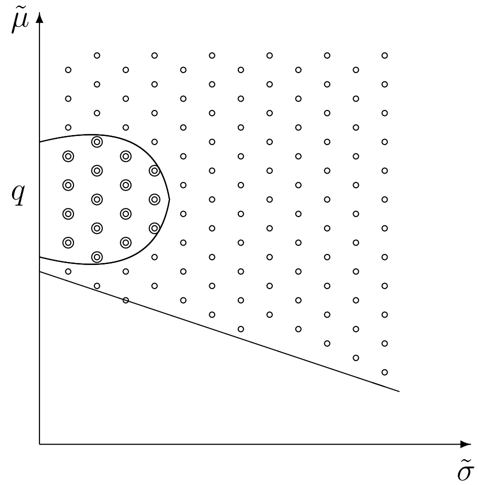

Figure 6: Optimal application policy of students as a function of posterior mean and standard deviation when there is noise in the decision process in the case when $c \leq \frac{\eta}{4}$.

References

Anderson, A. and L. Smith (2002). Assortative matching, reputation and the beatles breakup. University of Michigan, Department of Economics, unpublished.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy 81*(4), 813–846.

Brewer, D. J., E. R. Eide, and R. G. Ehrenberg (1999, Winter). Does it pay to attend an elite private college? cross-cohort evidence on the effects of college type on earnings. *Journal of Human Resources 34*(1), 104–123.

Burdett, K. and M. Coles (1997). Marriage and class. *Quarterly Economic Review 112*(1), 141–168.

Burdett, K. and M. Coles (1999, June). Long-term partnership formation: Marriage and employment. *Economic Journal 109*(456), F307–F334.

Gale, D. and L. Shapley (1962). College admissions and the stability of marriage. *American*

Mathematical Monthly 69, 9–15.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory* (2 ed.). The MIT Press.

Shimer, R. (2001). The assignment of workers to jobs in an economy with coordination frictions. NBER Working Paper, 8501.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

A Appendix

Proof of Theorem 3. The payoff from applying to school $\bar{q}$ is

$$\bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)\eta(\bar{q})-c.$$

This payoff is increasing in $\tilde{\mu}$, hence at any given $\tilde{\sigma}$ for a high enough value of $\tilde{\mu}$, a student with belief $F(\tilde{\mu},\tilde{\sigma})$ will prefer to apply to school i than not to apply at all as long as $\eta(\bar{q})>c$, while for low enough value of $\tilde{\mu}$, a student with belief $F(\tilde{\mu},\tilde{\sigma})$ will prefer to not apply at all to college to applying to school $\bar{q}$.

The difference in payoff from applying to the best school $\bar{q}$ and to a lower quality school $\hat{q}<\bar{q}$ is

$$\begin{aligned}
&\bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)\eta(\bar{q})-\bar{F}\left(\frac{\hat{q}-\tilde{\mu}}{\tilde{\sigma}}\right)\eta(\hat{q})= \\
&= \bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)[\eta(\bar{q})-\eta(\hat{q})]-\left[\bar{F}\left(\frac{\hat{q}-\tilde{\mu}}{\tilde{\sigma}}\right)-\bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)\right]\eta(\hat{q}) \\
&\geq \bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)[\eta(\bar{q})-\eta(\hat{q})]-\left[\bar{F}\left(\frac{\hat{q}-\bar{q}}{2\tilde{\sigma}}\right)-\bar{F}\left(\frac{\bar{q}-\hat{q}}{2\tilde{\sigma}}\right)\right]\eta(\hat{q}),
\end{aligned}$$

where the inequality follows from the fact that, for a given a, $F(x+a)-F(x)$ is maximized at $x=-a/2$. Notice that since $\bar{F}\left(\frac{\hat{q}-\bar{q}}{2\tilde{\sigma}}\right)-\bar{F}\left(\frac{\bar{q}-\hat{q}}{2\tilde{\sigma}}\right)$ is decreasing in $\tilde{\sigma}$ with $\lim_{\tilde{\sigma}\to\infty}\bar{F}\left(\frac{\hat{q}-\bar{q}}{2\tilde{\sigma}}\right)-\bar{F}\left(\frac{\bar{q}-\hat{q}}{2\tilde{\sigma}}\right)=0$ and $\bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)$ is increasing in $\tilde{\sigma}$ with $\lim_{\tilde{\sigma}\to\infty}\bar{F}\left(\frac{\bar{q}-\tilde{\mu}}{\tilde{\sigma}}\right)=\bar{F}(0)$, the lower bound on the difference is increasing in $\tilde{\sigma}$ with a strictly positive limit as $\tilde{\sigma}\to\infty$, hence there exists $\hat{\sigma}<\infty$ such that for all $\tilde{\sigma}>\hat{\sigma}$ it is preferable for a student to apply only to the best school than to apply only to one of the other schools. Depending then on whether the value from applying to the best school is positive or not, the acceptance set for any $\tilde{\mu}\in\mathbf{R}$ is $A(\tilde{\mu},\tilde{\sigma})=\{\bar{q}\}$ or $A(\tilde{\mu},\tilde{\sigma})=\emptyset$. □

Proof of Proposition 4. Notice that $P_1(\tilde{\mu},\tilde{\sigma})$ can be rewritten as

$$P_1(\tilde{\mu},\tilde{\sigma})=\int_{-\infty}^{\infty}f(\nu)\,\bar{F}\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)d\nu.$$

Then

$$\frac{dP_1(\tilde{\mu},\tilde{\sigma})}{d\tilde{\mu}} = \frac{1}{\tilde{\sigma}}\int_{-\infty}^{\infty} f\left(\nu\right) f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right) d\nu \geq 0.$$

To determine the sign of $\frac{dP_1(\tilde{\mu},\tilde{\sigma})}{d\tilde{\sigma}}$, notice that

$$\begin{aligned}
\frac{dP_1(\tilde{\mu},\tilde{\sigma})}{d\tilde{\sigma}} &= \int_{-\infty}^{\infty} f\left(\nu\right) f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}^2} d\nu = \\
&= \int_{-\infty}^{\underline{\mu}-\tilde{\mu}} f\left(\nu\right) f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}^2} d\nu + \int_{\underline{\mu}-\tilde{\mu}}^{\infty} f\left(\nu\right) f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}^2} d\nu = \\
&= \int_{\underline{\mu}-\tilde{\mu}}^{\infty} \left(f\left(\nu\right) - f(2(\underline{\mu}-\tilde{\mu})-\nu)\right) f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}^2} d\nu, \qquad \text{(A-1)}
\end{aligned}$$

where the second equality follows from the fact that $f\left(\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}}\right)\frac{\underline{\mu}-\tilde{\mu}-\nu}{\tilde{\sigma}^2} = -f\left(\frac{\nu-(\underline{\mu}-\tilde{\mu})}{\tilde{\sigma}}\right)\frac{\nu-(\underline{\mu}-\tilde{\mu})}{\tilde{\sigma}^2}$. Notice that $\underline{\mu}-\tilde{\mu}-\nu$ is always negative over the interval $[q-\tilde{\mu},\infty)$, and if $\tilde{\mu}\leq q$ then $f\left(\nu\right) - f(2(\underline{\mu}-\tilde{\mu})-\nu)$ is negative over the same interval, otherwise it is positive. This establishes the second part of Proposition 4. $\square$

Proof of Proposition 5. Notice that

$$\begin{aligned}
&\frac{d(P_2(\tilde{\mu},\tilde{\sigma})-P_1(\tilde{\mu},\tilde{\sigma}))}{d\tilde{\mu}} = -\frac{1}{\tilde{\sigma}^2}\int_{-\infty}^{\infty} f'\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right) F\left(q-\mu\right)\bar{F}\left(q-\mu\right) d\mu = \\
&= -\frac{1}{\tilde{\sigma}^2}\int_{\tilde{\mu}}^{\infty} f'\left[F\left(q-\mu\right)\bar{F}\left(q-\mu\right) - F\left(q-(2\tilde{\mu}-\mu)\right)\bar{F}\left(q-(2\tilde{\mu}-\mu)\right)\right] d\mu = \\
&= -\frac{1}{\tilde{\sigma}^2}\int_{\tilde{\mu}}^{\infty} f'\left[F\left(q-\mu+2(\mu-\tilde{\mu})\right) - F\left(q-\mu\right)\right]\left[F\left(\mu-q+2(q-\tilde{\mu})\right) - F\left(\mu-q\right)\right] d\mu,
\end{aligned}$$

where the argument of f' is suppressed and where the second equality follows from the fact that $f'(x) = -f'(-x)$, and the third equality follows from $x(1-x)-y(1-y) = (y-x)(y-(1-x))$. Clearly, for $\mu \geq \tilde{\mu}$, the second term in the integrand is always positive, while the third term is positive if and only if $\tilde{\mu}\leq q$. Hence $\frac{d(P_2(\tilde{\mu},\tilde{\sigma})-P_1(\tilde{\mu},\tilde{\sigma}))}{d\tilde{\mu}}$ is positive if and only if $\tilde{\mu}\leq q$.

To determine the sign of $\frac{d(P_2(\tilde{\mu},\tilde{\sigma})-P_1(\tilde{\mu},\tilde{\sigma}))}{d\tilde{\sigma}}$, notice that

$$\begin{aligned}\frac{d(P_2(\tilde{\mu},\tilde{\sigma})-P_1(\tilde{\mu},\tilde{\sigma}))}{d\tilde{\sigma}} &= -\frac{1}{\tilde{\sigma}^2}\int_{-\infty}^{\infty}\left[f+f'\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right]F\left(q-\mu\right)\bar{F}\left(q-\mu\right)d\mu = \\ &= \frac{1}{\tilde{\sigma}^2}\int_{-\infty}^{\infty}f\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]F\left(q-\mu\right)\bar{F}\left(q-\mu\right)d\mu,\end{aligned}$$

where once again the arguments of f and f' are suppressed, and where the second equality follows from the property of the normal distribution that $f'(x) = -f(x)x$. Since $f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]$ is symmetric around $\tilde{\mu}$, takes on negative values on $[\tilde{\mu}-\tilde{\sigma},\tilde{\mu}+\tilde{\sigma}]$ and positive values otherwise, and given $\int_{-\infty}^{\infty}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]d\mu = 0$, we know that

$$\int_{\tilde{\mu}}^{\tilde{\mu}+\tilde{\sigma}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}\right]d\mu = \int_{\tilde{\mu}+\tilde{\sigma}}^{\infty}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]d\mu,$$

and

$$\int_{\tilde{\mu}-\tilde{\sigma}}^{\tilde{\mu}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}\right]d\mu = \int_{-\infty}^{\tilde{\mu}-\tilde{\sigma}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]d\mu.$$

Now consider the case when $\tilde{\mu} \leq q$. Then $F\left(\frac{q-\mu}{\sigma_\nu}\right)\bar{F}\left(\frac{q-\mu}{\sigma_\nu}\right)$ is increasing over the interval $(-\infty,\tilde{\mu}]$, hence

$$\begin{aligned}&\int_{-\infty}^{\tilde{\mu}-\tilde{\sigma}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]F\left(\frac{q-\mu}{\sigma_\nu}\right)\bar{F}\left(\frac{q-\mu}{\sigma_\nu}\right)d\mu \leq \\ &\leq \int_{-\infty}^{\tilde{\mu}-\tilde{\sigma}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]F\left(\frac{q-(\tilde{\mu}-\tilde{\sigma})}{\sigma_\nu}\right)\bar{F}\left(\frac{q-(\tilde{\mu}-\tilde{\sigma})}{\sigma_\nu}\right)d\mu = \\ &= \int_{\tilde{\mu}-\tilde{\sigma}}^{\tilde{\mu}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}\right]F\left(\frac{q-(\tilde{\mu}-\tilde{\sigma})}{\sigma_\nu}\right)\bar{F}\left(\frac{q-(\tilde{\mu}-\tilde{\sigma})}{\sigma_\nu}\right)d\mu \leq \\ &\leq \int_{\tilde{\mu}-\tilde{\sigma}}^{\tilde{\mu}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[1-\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}\right]F\left(\frac{q-\mu}{\sigma_\nu}\right)\bar{F}\left(\frac{q-\mu}{\sigma_\nu}\right)d\mu,\end{aligned}$$

which implies that

$$\int_{-\infty}^{\tilde{\mu}}f\left(\frac{\mu-\tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu-\tilde{\mu})^2}{\tilde{\sigma}^2}-1\right]F\left(\frac{q-\mu}{\sigma_\nu}\right)\bar{F}\left(\frac{q-\mu}{\sigma_\nu}\right)d\mu \leq 0. \tag{A-2}$$

Similarly, if $\tilde{\sigma} \geq 2(q-\tilde{\mu})$, then $F\left(\frac{q-\mu}{\sigma_\nu}\right)\bar{F}\left(\frac{q-\mu}{\sigma_\nu}\right) \geq F\left(\frac{q-(\tilde{\mu}+\tilde{\sigma}))}{\sigma_\nu}\right)\bar{F}\left(\frac{q-(\tilde{\mu}+\tilde{\sigma})}{\sigma_\nu}\right)$ for any $\mu \in$

$[\tilde{\mu}, \tilde{\mu} + \tilde{\sigma}]$, and hence

$$\int_{\tilde{\mu}}^{\tilde{\mu}+\tilde{\sigma}} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right)\left[1 - \frac{(\mu - \tilde{\mu})^2}{\tilde{\sigma}^2}\right] F\left(\frac{q - \mu}{\sigma_\nu}\right) \bar{F}\left(\frac{q - \mu}{\sigma_\nu}\right) d\mu \geq$$

$$\geq \int_{\tilde{\mu}+\tilde{\sigma}}^{\infty} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu - \tilde{\mu})^2}{\tilde{\sigma}^2} - 1\right] F\left(\frac{q - \mu}{\sigma_\nu}\right) \bar{F}\left(\frac{q - \mu}{\sigma_\nu}\right) d\mu,$$

which implies that

$$\int_{\tilde{\mu}}^{\infty} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu - \tilde{\mu})^2}{\tilde{\sigma}^2} - 1\right] F\left(\frac{q - \mu}{\sigma_\nu}\right) \bar{F}\left(\frac{q - \mu}{\sigma_\nu}\right) d\mu \leq 0,$$

which together with Equation (A-2) means that

$$\int_{-\infty}^{\infty} f\left(\frac{\mu - \tilde{\mu}}{\tilde{\sigma}}\right)\left[\frac{(\mu - \tilde{\mu})^2}{\tilde{\sigma}^2} - 1\right] F\left(\frac{q - \mu}{\sigma_\nu}\right) \bar{F}\left(\frac{q - \mu}{\sigma_\nu}\right) d\mu \leq 0,$$

establishing the second part of Proposition 5. $\square$