Specialization in Higher Education*

Sinan Sarpca†

Koç University

October 2, 2007

*I am indebted to Dennis Epple and Holger Sieg for their constant support. I would like to thank Richard Romano, Eric Hanushek, Marty Gaynor, Ron Goettler, Maria Ferreyra, Michael Peress, Insan Tunali, Bruce Weinberg, Kevin Lang, Fallaw Sowell, George-Levi Gayle, Limor Golan, Peter Arcidiacono, Stephen Spear, Kenneth Smith, Robert Toutkoushian, in addition to workshop participants at Carnegie Mellon, Koç University, EEA in Vienna, ESEM in Budapest, NASM at Duke, and APET in Nashville for their comments and suggestions. Any errors are mine. I also would like to thank Bill Elliott, Vice President for Enrollment at Carnegie Mellon, for the data used in this paper.

†E-mail: ssarpca@ku.edu.tr; Phone: +90 (212) 338-1214; Address: Koç University, Rumelifeneri Yolu, Sariyer 34450 Istanbul
Abstract

This paper develops a differentiated products model of school competition that distinguishes among different dimensions that matter in the skill acquisition process. The model predicts that when schools compete for students, specialization may arise as a competition strategy. This will serve rich students well. Poorer students, however, may attend schools with specializations that do not cater to their relative talents. By doing so, these poorer students complement the weaknesses of the richer students through peer effects and receive financial aid in return. The empirical analysis provides strong support for the model’s predictions about within-school implications of specialization.

Keywords: higher education, school competition, peer effects, specialization, multi-dimensional skills.

JEL Classification: I21, D58.
1 Introduction

Research in labor economics provides plenty of evidence that skills are multi-dimensional. However, almost every paper with peer effects in education assumes that the quality of a peer group can be captured by a single dimensional index. The models that adopt this assumption also fail to explain specialization of schools (e.g., technical universities, liberal arts colleges, elementary and secondary magnet schools), which is another observed fact. To complement the research on school competition, peer effects, and related areas, there are large potential benefits of studying a differentiated products model that properly distinguishes among different dimensions that matter in the skill acquisition process.

Students differ in mathematical, verbal, and other capabilities. Schools often specialize to serve students with different capabilities and interests. Peer interactions may be key to understanding such specialization. A given student’s achievement may well be affected by the overall ability of the peers in the school the student attends. However, the effect of the peer group surely depends on the extent to which peers’ abilities and interests are congruent with a student’s own abilities and interests. Specialization makes a college more appealing to students who have relatively strong skills in the area in which it specializes. However, while attempting to graduate students who are exceptionally strong in its area of specialization, a college may prefer to avoid the pitfall of providing an overly narrow education. Hence, a college will generally offer education in subjects outside its area of specialization, creating the challenge of attracting able students that are interested in those subjects. Financial aid is one tool that a college may attempt to use to attract strong students to its weaker departments.

In the first two parts of the paper, I provide a formal economic model and derive predictions about financial aid policies and student composition of a college. I consider, in particular, the implications of two dimensions of ability, with the intention of providing predictions that may be tested using SAT scores that measure students’ mathematical and verbal aptitudes. This model predicts that able high-income students will attend colleges with specializations congruent with their interests and receive little financial aid. By contrast, able low-income students will attend schools with specializations opposite to their relative
talents, complement the weaknesses of the wealthier students through peer effects, and receive financial aid in return. In the third part of the paper, I conduct an empirical analysis using data on the entire set of applicants of Carnegie Mellon University. The empirical evidence provides strong support for these financial aid predictions, and suggests that lower income students’ educational choices are likely to be affected by schools’ financial aid policies. These findings are of potential importance to government policy makers in framing financial aid policies to assist lower-income students. Policy makers may wish to design financial aid policies that enable poorer students to attend colleges better suited to their relative talents.

This paper contributes to research related to peer effects by considering multidimensional effects that operate non-uniformly through interaction with an individual’s own characteristics. It contributes to research on school competition by considering separate treatment of different types of ability, within-school variation of quality, associated effects on financial aid, and the equilibrium consequences to students of differing incomes and abilities. Although the particular focus is on four-year degree granting colleges and universities, this analysis can also be used to address questions related to specialization in earlier stages of education such as magnet schools, charter schools, specializing prep academies, etc.

The existence of peer effects in primary and secondary education has been discussed since the Coleman Report (1966). Various reform proposals led researchers to analyze market implications of school choice in the presence of peer effects (e.g., Epple and Romano (1998), Nechyba (2000)). More recently the discussion turned to higher education and peer effects have been documented by Zimmerman (2003), Sacerdote (2001), Stinebrickner and Stinebrickner (2006), and Winston and Zimmerman (2004). The predictions that I derive and test arise either if there are peer effects or preferences for attending a college with high-ability peers. As long as students and parents value peer quality, and schools price it accordingly,

---


2The empirical evidence on peer effects is controversial. Several recent studies suggest that the commonly used specifications may be one reason. Dale and Krueger (2002) document a relation between dispersion of ability in a college and earnings of graduates. This suggests that students may care about more than just the average level of ability in a school. Hoxby and Weingarth (2005) provide a criticism of the commonly used linear-in-means and single-crossing models. The multi-dimensional non-uniform peer effects specification in this paper provides an alternative approach.
the market implications will be identical. Previous work on college choice and pricing includes Manski and Wise (1983), Ehrenberg and Sherman (1984), Rothschild and White (1995), Dale and Krueger (2002), and Hoxby (2004), among others. Epple, Romano, and Sieg (2003, 2006) also construct a formal model of college competition in the presence of peer effects, and show that observed financial aid policies and student allocations in U.S. are consistent with their model’s predictions. While providing some complementary analysis, this paper tackles a different question -specialization- and the model in this paper differs in treatment of student abilities, heterogeneity of students’ experiences of quality, and school qualities and objectives. In the empirical part, this paper focuses on within-school implications.

The rest of the paper is organized as follows: Sections 2 and 3 describe the theoretical framework and provide a number of predictions about equilibrium implications. The data are discussed in Section 4. Section 5 presents the empirical findings. Section 6 concludes.

2 The Model

A college application package documents a prospective student’s various skills and achievements in different subjects. Mathematical and verbal skills are demonstrated by standardized test scores such as SAT, or SAT II Writing and Math subject tests. Other skills are demonstrated through instruments that are harder to standardize and convert into numbers, such as recommendation letters and personal essays. Also, many students submit a financial aid application documenting their ability to pay for tuition and expenses. In the model, an applicant’s different abilities are objectively measured and summarized in a vector $b = (b^m, b^v)$.\(^3\)

Income is denoted by $y$ and is known to every college. The vector $(b^m, b^v, y)$ characterizes a student. These student characteristics have the joint marginal distribution $f(b^m, b^v, y)$, continuous and positive on its support $S \equiv (0, b_{\text{max}}]^2 \times (0, y_{\text{max}}]$.

An applicant has the utility function $U(\cdot)$ increasing in its two components: numeraire consumption and

\(^3\) The vector $b$ is two dimensional mainly for expositional convenience. First four propositions can immediately be generalized for $n$-dimensional vectors.
educational achievement. Consumption is equal to the income after tuition expenditure \( p \). Educational achievement is increasing in student’s own capabilities \( b = (b^m, b^v) \) as well as the attended school’s quality.

There are a finite number of competitive schools \((i = 1, \ldots, n)\). A school’s quality is determined by the quality of the peer group \( \Theta_i = (\theta^m_i, \theta^v_i) \), where each component represents the mean level of that type of ability in the school. So the utility of an applicant is given by \( U (y - p, a (\theta^m, \theta^v, b^m, b^v)) \) with \( U_1, U_2, \) and all four partial derivatives of \( a \) being positive. Students are free to not attend college. In that case, they pay no tuition \( (p_0 = 0) \), and experience a fixed quality \( \Theta_0 = (\theta^m_0, \theta^v_0) \).\(^4\) I assume \( \theta^m_0 \) equals \( \theta^v_0 \) and college qualities exceed those of the outside alternative, i.e., \( \min_{i,j \in \{1, \ldots, n\}} \{\theta^m_i, \theta^v_j\} \geq \max\{\theta^m_0, \theta^v_0\} \).\(^5\)

The function \( U (.) \) satisfies the single-crossing (in income) condition:

\[
\frac{\partial}{\partial y} \left( \frac{\partial U}{\partial a} \right) / \partial y > 0 \quad \text{(SCy)}
\]

This condition ensures that among two students with identical ability, the one with higher income is willing to pay more for an increase in achievement level.

All colleges have the same cost function:

\[
C (k) = V (k) + F
\]

where \( k \) denotes the size of the college, \( V' > 0 \) and \( V'' > 0 \), and \( k^* \) denotes the “efficient scale”:

\[
k^* = \arg \min_k \frac{C (k)}{k}
\]

Lack of such an economies of scale assumption implies an infinite number of schools in equilibrium.

The colleges maximize profits by choosing tuition and admission policies behaving as utility takers.\(^6\)

---

\(^4\)Alternatively, the quality of the outside option can be equal to the average math and verbal ability of those choosing it. Doing so does not alter the qualitative results in this paper.

\(^5\)This is for simplicity of exposition. Since I consider only desirable peer characteristics here, this could be obtained as an equilibrium property even if it was not assumed.

\(^6\)See Scotchmer (1994). An alternative objective used in the literature is quality maximization (Epple, Romano, and Sieg,
Utility taking is a generalized version of price taking when consumers and products differ; when choosing
tuition and admission policies, a college takes its competitors’ policies and resulting student utility levels
as given. There is free entry and exit. Let \( \alpha_i(b^m, b^v, y) \) denote the proportion of the type \((b^m, b^v, y)\) that
college \( i \) admits. The college’s problem can be written as:

\[
\max_{\theta^m_i, \theta^v_i, k_i, \alpha_i(b^m, b^v, y), p_i(b^m, b^v, y)} \int \int \int_S p_i(b^m, b^v, y) \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy - V(k_i) - F
\]

subject to the constraints:

\[
U(y - p_i(b^m, b^v, y), a(\theta^m_i, \theta^v_i, b^m, b^v)) \geq \max_{j \in \{0,1,\ldots,n\}} U(y - p_j(b^m, b^v, y), a(\theta^m_j, \theta^v_j, b^m, b^v)) \quad \forall (b^m, b^v, y);
\]

\[
\alpha_i(b^m, b^v, y) \in [0, 1] \quad \forall (b^m, b^v, y);
\]

\[
k_i = \int \int \int_S \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy;
\]

\[
\theta^m_i = \frac{1}{k_i} \int \int \int_S b^m \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy;
\]

\[
\theta^v_i = \frac{1}{k_i} \int \int \int_S b^v \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy;
\]

Constraint (3) imposes the utility taking assumption. To be able to admit a student, a college must provide
a utility level the student can achieve by choosing any other alternative. Constraint (4) guarantees that
schools will admit a nonnegative proportion of a type if they admit any, and also will not admit more
than what is available. Given the admitted proportions of each type, constraints (5) and (6) calculate the
college’s size and quality vector.

(2003), (2006)). I discuss below that a single dimensional quality measure for schools is not appropriate for this setup.
The equilibrium is described by the following conditions:

\[
\forall (b^m, b^v, y), \quad U^* (b^m, b^v, y) = \max_{i \in \{0, 1, \ldots, n | \alpha_i (b^m, b^v, y) > 0 \text{ is in the optimal set of } i\}} U (y - p_i (b^m, b^v, y), a (\theta_i^m, \theta_i^v, b^m, b^v)) \quad \text{(UM)}
\]

\[
[\theta_i^m, \theta_i^v, k_i, p_i (b^m, b^v, y), \alpha_i (b^m, b^v, y)] \text{ satisfy (2) – (6) for } i = 1, 2, \ldots, n \quad \text{(IIIM)}
\]

\[
\pi_i \equiv \int \int \int_S p_i (b^m, b^v, y) \alpha_i (b^m, b^v, y) f (b^m, b^v, y) db^m db^v dy - V (k_i) - F = 0 \quad i = 1, 2, \ldots, n \quad \text{(ZIII)}
\]

\[
p_o (b^m, b^v, y) = 0 \text{ for all } (b^m, b^v, y), \quad \text{(NC)}
\]

\[
\alpha_o (b^m, b^v, y) \in [0, 1] \text{ for all } (b^m, b^v, y),
\]

\[
k_o = \int \int \int_S \alpha_o (b^m, b^v, y) f (b^m, b^v, y) db^m db^v dy;
\]

\[
\sum_{i=0}^n \alpha_i (b^m, b^v, y) = 1 \quad \forall (b^m, b^v, y) \quad \text{(MC)}
\]

Condition (UM) describes students’ utility maximization. A student can attend one of the colleges that will admit him/her, or simply may choose not to attend college. Among these, the student will choose a utility maximizing alternative taking pricing and admissions policies as given. Conditions (IIIM) and (ZIII) summarize profit maximization and free entry-exit assumptions of the model. Condition (NC) describes the no-college option. It is free and is open to everyone. Finally condition (MC) implies market clearance: every student will either be enrolled in a school or be matched to the no-college alternative in equilibrium.
3 Theoretical Results

3.1 Properties of Equilibrium

The first-order conditions for the school’s problem described in equations (2) through (6) are as follows:

\[ U (y - p_i^*, a(\Theta_i, b)) = U^* (b^m, b^v, y) \]  

\[ \alpha_i \begin{cases} 
  = 0 \\
  \in [0, 1] \\
  = 1
\end{cases} \quad \alpha_i \begin{cases} 
  \alpha_i \begin{cases} 
  < \quad \text{as } p_i^* (b^m, b^v, y, \theta_{m_i}, \theta_{v_i}) \\
  > \quad \frac{V' (k_i) + \eta_{im} (\theta_{m_i} - b^m) + \eta_{iv} (\theta_{v_i} - b^v)}{EMC_{(b^m, b^v)}}
\end{cases}
\end{cases} \]  

\[ \eta_{im} = \frac{1}{k_i} \int \int \int_S \frac{\partial p_i^*}{\partial \theta_{m_i}} \alpha_i f \, db^m \, db^v \, dy \]  

\[ \eta_{iv} = \frac{1}{k_i} \int \int \int_S \frac{\partial p_i^*}{\partial \theta_{v_i}} \alpha_i f \, db^m \, db^v \, dy \]  

Condition (7) describes a school’s optimal tuition policy. The RHS term is the student’s reservation utility as described in condition (UM). The optimal tuition is the one that makes the student indifferent to his/her best alternative. This is the maximum tuition that the college can charge and still attract the student. Suppose the college admits a student with say mathematical ability that is below the average of the school. This will decrease the math quality (as defined by average math ability) and will result in a lower willingness to pay by all other students. The opposite argument will hold with a student above the average. The Lagrange multipliers in (9) calculate these marginal effects deriving from changes in \( \theta_{m_i} \) and \( \theta_{v_i} \). Consider the RHS term in (8). The first component is the marginal cost that comes from cost function in (1). The second and third terms represent the marginal cost arising from the student’s (negative or positive) contribution on peer quality. For example if a student’s ability \( b^m \) is below the mean ability \( \theta^m \), the student pays \( \eta_{im} \times (\theta_{m_i} - b^m) \) to compensate for the negative contribution, whereas if its above \( \theta^m \),
the student receives the same amount in tuition discounts for increasing peer quality. This RHS term is
the marginal cost function that internalizes the externalities and we call it the effective marginal cost of a
student. Condition (8) describes the optimal admission policies. It states that a school will admit students
with reservation prices that equal or exceed effective marginal cost only.\footnote{The first-order conditions are sufficient for a maximum when \( \alpha_i(b^m, b^r, y) = 0 \) or 1. When \( \alpha_i(b^m, b^r, y) \in (0, 1) \), a cost function with a large curvature (high \( V'' \)) is sufficient for a local maximum. A proof is available upon request from the author.}

The implications of the conditions (7)-(9) can be summarized as follows:

**Proposition 1** No two colleges will provide an identical quality profile (i.e., \( \theta_i^m = \theta_j^m \Rightarrow \theta_i^v \neq \theta_j^v \) for all \( i, j \in \{0, 1, ..., n\} \) s.t. \( i \neq j \)).

The proof is in the appendix. If two colleges are exactly identical in terms of quality, they would engage
in price competition for every student, charging a tuition equal to effective marginal cost. If one college
differentiates itself at least in one dimension, it can gain room for pricing according to willingness to pay for
quality. When quality is measured on a single dimension, the above proposition implies a strict hierarchy
of school qualities (as in Epple and Romano (1998)). In multiple dimensions however, schools may also
differentiate themselves by specialization.

Let \( A_i \equiv \{(b^m, b^r, y) \in S|\alpha_i(b^m, b^r, y) > 0 \text{ is optimal when } p_i(b^m, b^r, y) = EMC_i(b^m, b^r)\} \) denote the
admission space of school i. A boundary locus between schools i and j is \( A_i \cap A_j \), if it exists. The pricing
implications are summarized in the following result:

**Proposition 2** (i) On a boundary locus between schools i and j, tuition equals effective marginal cost of
a student at either school, thus depends only on ability. (ii) Off-boundary students are charged tuitions
greater than effective marginal costs, depending on income as well as ability. (iii) In equilibrium, every
student attends one of the schools that would maximize utility if all schools instead set tuition equal to
equilibrium effective marginal cost for all students.
A student on the boundary locus between two schools is indifferent to attending either one of them at prices equal to effective marginal costs. A student on the interior of a school’s admission space however, has a strict preference for that school if priced at effective marginal cost. Schools exploit this by increasing tuition up to the point where the student becomes indifferent between attended school and second best. The presence of close-competitor schools would limit this price discrimination. Then the tuition will be approximately equal to the EMC term in (8), and will vary with each dimension of student ability, and the quality of school on each dimension.

**Proposition 3** If \((b^m, b^v, y_1)\) chooses school \(i\) and \((b^m, b^v, y_2)\) chooses school \(j\) where \(a(\theta^m_i, \theta^v_i, b^m, b^v) \leq a(\theta^m_j, \theta^v_j, b^m, b^v)\) then \(y_2 \geq y_1\).

This follows from the single crossing property of preferences and is straightforward. Proposition 3 gives a partial picture of the stratification in equilibrium. For every group of students with identical ability vector \((b^m, b^v)\), there are income thresholds separating schools’ admission spaces, with higher income students attending higher achievement schools for that ability type.

Development of additional implications is facilitated by imposing additional structure in the model. That requires further specification. Adopting a single-dimensional quality measure in this framework simplifies the analysis and strengthens the results, but proves to be overly restrictive, as I now demonstrate. Consider a quality function \(q_i = q(\theta^m_i, \theta^v_i)\) (e.g., a weighted average), where achievement of a student depends on \(q\) but not on individual \(\theta\)'s, i.e., \(a(q, b^m, b^v)\). With one quality dimension, a strict hierarchy of schools is implied:

**Proposition 4** When achievement is a function of \(q = q(\theta^m, \theta^v)\), a strict hierarchy of school qualities result: \(q_n > q_{n-1} > \ldots > q_1\)

This is a direct corollary of Proposition 1.
3.2 Specialization

As the preceding discussion demonstrates, analysis of specialization requires potential for richer interaction between a student’s own abilities and peer abilities. Hence, I propose and use the additively separable achievement function:

\[
a(\Theta, b) = \left[ (\theta_m)^\beta (b_m)^\gamma + (\theta_v)^\beta (b_v)^\gamma \right]^{\frac{1}{\delta}}
\]

\[\delta > 0; \quad \beta, \gamma \in [0, 1].\]  

(10)

This specification implies that a unidimensional measure for college quality is not appropriate for everyone, and the ranking of colleges according to potential achievement depends on a student’s own abilities as well. The students’ preferences over colleges are determined by the interaction, or fit, between student and college characteristics. This feature of the model makes strong predictions about financial aid policies of colleges and resulting within college variation of students, neither of which are captured in the one-dimensional models.

To focus on the effects of competition by specialization, I consider an economy with two colleges (i=1,2), no-college option (i=0), and a student population with symmetrically distributed verbal and mathematical abilities, i.e., \(f(b^m, b^v, y) = f(b^v, b^m, y)\). The utility function of a household is \(U = (y - p) a(.)\). I consider an allocation that satisfies the following five properties: a) market clearing; b) utility maximization; c) consistency of reservation utilities; 8 d) the first-order conditions of the colleges’ optimization problems; and e) the second-order conditions are satisfied locally. An allocation satisfying these properties is termed a “local equilibrium” by Epple, Romano, and Sieg (2006). 9 Given the symmetry of the distribution of types in the population, it is natural to consider the local equilibrium in which schools are symmetric, i.e., \(\theta_m^1 = \theta_v^2, \text{ and } \theta_v^1 = \theta_m^2\).

---

8This refers to the property that every school takes the utility defined in (UM) as the reservation utility in equilibrium. 9As explained in Epple, Romano, and Sieg (2006), global equilibrium requires that the allocation is globally optimal for each college. Global optimality can be verified computationally once a local allocation has been found. In the computations, conditions (a) through (d) are used to calculate a candidate local equilibrium. It is then verified computationally that (e) holds.
Given an ability profile \((b^m, b^v)\), Proposition 3 implies that there are income thresholds for attending first best college, attending second best college, and no college. Given qualities \(\Theta_i = (\theta^m_i, \theta^v_i)\) and \(EMC_i (b^m, b^v)\) for options \(i = 0, 1, 2\), these thresholds \(\{y^*_{ij} (b^m, b^v): i, j \in \{0, 1, 2\}, i \neq j\}\) can be mapped by an investigation of boundary students using Proposition 2 (see Appendix for derivations). Consider the following partition of the student type space \(S\) with respect to relative ability and income thresholds described above:

\[
M^h = \left\{(b^m, b^v, y): b^m > b^v \text{ and } y > \max\{y^*_{12} (b^m, b^v), y^*_{01} (b^m, b^v)\}\right\}
\]

\[
M^l = \left\{(b^m, b^v, y): b^m > b^v \text{ and } y^*_{02} (b^m, b^v) < y < y^*_{12} (b^m, b^v)\right\}
\]

\[
V^h = \left\{(b^m, b^v, y): b^m < b^v \text{ and } y > \max\{y^*_{12} (b^m, b^v), y^*_{02} (b^m, b^v)\}\right\}
\]

\[
V^l = \left\{(b^m, b^v, y): b^m < b^v \text{ and } y^*_{01} (b^m, b^v) < y < y^*_{12} (b^m, b^v)\right\}
\]

\[
NC = S \setminus (M^h \cup M^l \cup V^h \cup V^l)
\]

We get a specialization equilibrium where one college is more desirable for half of the population, other is more desirable for the other half, and college attendance increases in each type of ability and income. Also, the symmetry of the distribution of student abilities implies a symmetry in shadow prices \((\eta^m_i = \eta^v_j; i, j \in \{1, 2\}, i \neq j)\). The student composition in each college, however, depends on the relationship between the college qualities and the shadow prices of \(b^m\) and \(b^v\) in equilibrium:

**Proposition 5** College quality profiles \(\{(\theta^m_1, \theta^v_1), (\theta^m_2, \theta^v_2)\}\) are such that \(\theta^m_1 = \theta^v_2 > \theta^v_1 = \theta^m_2\). Depending on the model parameters shadow prices can be related in two ways: (i) If \(\eta^m_1 = \eta^v_2 < \eta^m_2 = \eta^v_1\), then admission spaces of the two colleges are \(A_1 = M^h \cup V^l\) and \(A_2 = V^h \cup M^l\). (ii) If \(\eta^m_1 = \eta^v_1 \geq \eta^m_2 = \eta^v_2\), then \(A_1 = M^h \cup M^l\) and \(A_2 = V^h \cup V^l\). In both cases, those in set \(NC\) do not attend college.

The proof is in the appendix. Proposition 3 implies that students with highest incomes, \(M^h\) and \(V^h\), attend College 1 (higher \(\theta^m\)) and College 2 (higher \(\theta^v\)) respectively. However, that is a partial picture, and depending on model parameters, the rest of the seats in each school may fill in two different ways. Let’s
ignore the no college option for a moment.

The first allocation described in the proposition is where we see separation by both (relative) ability and income: We see a mixing of relative ability types; \( V^l \) joins \( M^h \) in College 1, and \( M^l \) joins \( V^h \) in College 2. In other words the student body in each college will consist of two groups: 1. High income students that fit well their college and benefit most from the specialization. 2. Low income students that could get a better education in the other college. The interpretation is as follows: A student \((b^m, b^v, y)\) such that \(b^m > b^v\) has a tuition advantage at College 2 \((\eta^m_1 < \eta^m_2 \text{ and } \eta^v_1 > \eta^v_2)\), but can get higher educational achievement at College 1 \((\theta^m_1 > \theta^m_2 \text{ and } \theta^v_1 < \theta^v_2)\). So student’s decision depends on whether income is “high enough” to substitute achievement gain for loss from higher tuition. The choice-switching income threshold \(y_{12}^*\) is increasing in \(b^m\) since higher \(b^m\) provides a higher tuition discount in College 2. It is increasing in \(b^v\) too, since that means higher achievement in College 2. A symmetric argument applies for a student with \(b^v > b^m\). Figure 1 illustrates the typical shape of this \(y_{12}^*\) surface obtained using computational models.

At any given income level \(\bar{y}\), the locus of \((b^m, b^v)\) pairs that satisfies \(y_{12}^*(b^m, b^v) = \bar{y}\) is a convex curve when plotted on a \(b^v - b^m\) plane. Those to the right of this curve are below their \(y_{12}^*(b^m, b^v)\), and they attend the school with specialization opposite to their relative talent. Those to the left of the curve have higher income than their \(y_{12}^*\) and they attend to the higher-achievement school for them. The curve is convex because it is easier to substitute one school for another when \(b^m\) is close to \(b^v\) (one with \(b^m = b^v\) is indifferent between two schools, independent of school qualities). Note that some students will attend the opposite school although they have high incomes (For example, see NE corner in Panel 3 of Figure 2). There are two reasons for that. First, their \(b^m\) and \(b^v\) are very high compared to \(\theta^m\) and \(\theta^v\) at either school, they don’t benefit much from either peer groups. Second, their \(b^m\) and \(b^v\) are close so schools are easily substitutable. On the other hand, tuition advantage is greatest for those because of the very high ability levels. The attendance of low income students to the higher-achievement school can be understood by similar arguments.

The second possible allocation is one where we see stratification by (relative) ability, i.e., \(M^l\) will join
$M^h$ in College 1, and College 2 will consist of $V^h$ and $V^l$. In equilibrium everyone attends the college that is best for them, regardless of their income. Even though the entire peer group consists of people with similar skills, another similar student is more desirable than a student with disparate strengths. Even though it is not possible to establish analytically, computational models suggest that this is a degenerate equilibrium.

To see how the no-college option alters the picture, remember that allocation of students to schools is as if they are priced at effective marginal costs (Proposition 2). Then an inspection of the EMC function in (8) shows that tuition in each school decreases in each ability type. This means that holding income constant, school attendance increases in either ability. Also Proposition 3 implies that holding abilities constant, school attendance should increase in income. Therefore, a surface that starts at some high income intercept and decreases in each ability separates the part of the type space that choose no-college (See Figure 2). This concludes the discussion of Proposition 5.

I investigate the relation between relative abundance of an ability and its shadow price using computational models. This analysis also suggests a robust relation between college qualities and the extent of within-college variation of abilities. I report one such model below.

### 3.3 An Example

Both ability types are distributed normal with a mean of 500 and a standard deviation of 100 points, censored at 200 at the bottom and at 800 at the top. Suppose there is a 0.50 correlation between $b^m$ and $b^v$, and a 0.25 correlation between income and each ability. Also, the logarithm of income is distributed normal with mean 3.36 and standard deviation 0.68.

There are two schools, and an outside-college option at no cost with quality $\theta_o = \theta_o^m = \theta_o^v = 1$. Schools have the same quadratic cost function $C(k) = 150 + 10,000k + 1,389k^2$. I set $\beta = 0.2, \gamma = \delta = 1$. Those parameters are chosen so that: a) one third of the population attends college in equilibrium;\(^10\) b) Average

\(^{10}\)That is approximately the ratio attending four-year degree programs in U.S. (Epple et al.(2003)).
tuition is about $10,000; c) Equilibrium number of schools is two; d) A household’s utility increases by 10% if an average student’s ability increases by 10%, and by about 1.5% if the school’s quality increases by 10%. Equilibrium values are given in the table below. Sensitivity analysis shows different parameters give results qualitatively same. Admission spaces are depicted in Figure 2.

Example 1: Computed Equilibrium

<table>
<thead>
<tr>
<th>School</th>
<th>$\theta^m$</th>
<th>$\theta^v$</th>
<th>$\kappa$</th>
<th>$\eta^m$</th>
<th>$\eta^v$</th>
<th>$\sigma^m$</th>
<th>$\sigma^v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td>556</td>
<td>538</td>
<td>0.16</td>
<td>9.60</td>
<td>9.85</td>
<td>82</td>
<td>104</td>
</tr>
<tr>
<td>School 2</td>
<td>538</td>
<td>556</td>
<td>0.16</td>
<td>9.85</td>
<td>9.60</td>
<td>104</td>
<td>82</td>
</tr>
</tbody>
</table>

School 1 has a higher type-m quality, and provides higher achievement than school 2 to those with $b^m > b^v$. One point of $b^m$ is worth $9.60$ and one point of $b^v$ is worth $9.85$, i.e., rare ability has a higher shadow price. I found this relation to be robust to variations in model parameters. The reason we do not see a larger spread between $\theta^m$ and $\theta^v$ and one between $\eta^m$ and $\eta^v$ is partly due to school sizes in this example, and partly implied by the model: As $\theta^m$ increases in a school, $b^v$ becomes rare and $\eta^v$ increases. This attracts high $b^v$ students, and an increase in $\theta^v$ will follow. However, the difference between $\theta^m$ and $\theta^v$ and the one between $\eta^m$ and $\eta^v$ increase as school sizes decrease, as number of schools increase, or as the correlation between $b^m$ and $b^v$ decreases.

A second robust relation is the difference between variances of $b^m$ and $b^v$ in a school. The standard deviations are given in the last two columns of the above table. The rare ability has a higher standard deviation. The role of this property will be further discussed below.

### 3.4 Cross Subsidization

The stratification result of Epple and Romano (1998) implies a cross-subsidization in equilibrium. High-ability students with low incomes are subsidized by low ability-high income types via tuition discounts due to their contribution to mean ability. The model in this paper generates a cross-subsidization in a different form. Students with high incomes and relatively high ability of one type subsidize students with
low incomes and relatively high ability of the other type. For example, for students such that \( b^m > b^v \), College 1 provides higher achievement. But a considerable proportion of such students attend College 2. Although they sacrifice some potential achievement by switching to the lower \( \theta^m \) school: 1. they gain part of it back because of the higher \( \theta^v \), and: 2. are subsidized for the rest via tuition by high income - low \( b^m \) (and even lower \( b^v \)) type students.

My goal in the empirical analysis that follows will be documenting this phenomenon, in the light of the following three predictions of the theoretical model and computational analysis:11

Prediction 1. For two schools such that \( \theta^m_1 > \theta^m_2 \) and \( \theta^v_1 < \theta^v_2 \), the shadow prices of abilities are related as follows: \( \eta^m_1 < \eta^m_2 \) and \( \eta^v_1 > \eta^v_2 \). Above statement implies that a school will charge higher tuitions to students that fit better to its profile. Consider two students \( s \) and \( t \) with abilities \((b^m_s, b^v_s) = (b, b')\) and \((b^m_t, b^v_t) = (b', b)\) with \( b > b' \) and two schools such that \((\theta^m_1, \theta^v_1) = (\theta, \theta')\) and \((\theta^m_2, \theta^v_2) = (\theta', \theta)\) with \( \theta > \theta' \). Prediction 1 implies that, school-1 will offer more financial aid (or charge less tuition equivalently) to student \( t \) then it does to student \( s \).

Prediction 2. For two schools such that \( \theta^m_1 > \theta^m_2 \) and \( \theta^v_1 < \theta^v_2 \), the variances of abilities are related as follows: \( \sigma^m_1 < \sigma^m_2 \) and \( \sigma^v_1 > \sigma^v_2 \). As noted earlier, the above relationship is a robust feature of computed equilibria. This is caused by the pricing implication of Prediction 1. Such a pricing policy attracts students with particularly large (positive) deviations in the rare ability. The implication of admission spaces such as the one in Proposition 5 can be summarized as:

Prediction 3. The fit between student’s and school’s profile increases in income

Those three predictions characterize cross-subsidization. The first indicates the financial advantage of attending a school less desirable in terms of achievement. The second indicates that the rare ability has

\[11\text{The comparative statements are for schools with same sizes.}\]
a greater role in creating this advantage. The third completes the picture by indicating the direction of subsidies as from rich to the poor.

4 Data

To test my predictions, I use data on the entire set of applicants at Carnegie Mellon University in the year 2002. Carnegie Mellon is well-suited for testing the predictions of my model. It is a technical university by reputation. The relative abundance of math ability is evident in the SAT scores. It is highly selective and the peer group composition is carefully determined by admission policies. These are consistent with how a specializing school is defined in the model. The model is more appropriate for discussing private schools, since the objectives, financing, and the resulting admission and pricing policies of public universities may be more complicated. CMU is a private school.

The extensive data set includes scores on standardized tests, family income for those who applied for financial aid, financial aid awarded, the undergraduate area of specialization to which the student is admitted, demographic characteristics, and other information obtained during the admission process. Table 1 displays the descriptive statistics for my sample. CMU consists of six colleges. A candidate can apply to many colleges at once, and each application is treated independently. The main file I use contains 8983 applications from 8344 applicants, 4997 of which submit financial aid applications as well. Financial aid is offered to 2411 of 4392 accepted, and 1358 students enroll with 886 receiving aid.

When describing the model, I had mentioned the importance of non-academic skills that are signalled through personal essays, recommendation letters, or other forms and questionnaires in the application package. At CMU, admissions officers go through those documents and assign a number ranging from 0 to 4. That is called non-academic factor, and it is the main indicator of non-academic skills that are valuable from CMU’s standpoint. That allows me to incorporate a large amount of crucial information into the analysis as a standardized summary variable.

As academic ability measures, the data set contains standardized test scores and high school achievement
variables. I use SAT math and SAT verbal scores as main indicators of math and verbal abilities. For the enrolling class, the math scores have a mean of 717 and standard deviation of 61. The verbal scores have a mean 645 and standard deviation 81. As alternative measures, the subject tests SAT II writing test, and SAT II 1C and 2C (math) test scores are available as well. Four departments in the College of Fine Arts do not require SAT scores for admission, so I exclude those from my analysis.

Income related information is as crucial as ability variables for the analysis. The data set contains various income and need measures for those applied to financial aid. Need is the difference between the ability to pay and the total annual cost of education. I use a need measure calculated by CMU according to detailed information provided by applicants. Financial aid information is available in full detail. Some students negotiate for a better offer if they receive offers from competitor schools. If CMU responds, the increment is called a reaction gift. I discuss this in greater detail in the next section. As aid variables I use the total gift amount, both the initial offer, and final offer including the reaction award. The calculated need for the enrolling class ranges from high negative numbers (less than -$100,000) up to the full cost. Mean need is $9,647 and standard deviation is $37,070.

The gender and race composition is as follows: About 36 percent of the enrolling class is female, 7 percent classify themselves as African-American or Native-American, and 5 percent classify themselves as Hispanic. The composition of aid recipients is quite similar. Also, there are many other variables that I used to construct indicator variables to control for various issues. I do not report summary statistics for the ones I found to be insignificant. I discuss some of those variables in the following section.

5 Empirical Analysis

5.1 Student Composition:

According to the model, higher income students in a school fit better to the school’s profile. Lower income students could get a better education at another school with a profile that fits them better, but are attracted
to this school because of the tuition discounts. These discounts are subsidized by higher-income/better-fitting students through the pricing mechanism. At CMU, that means their math skills are stronger, and verbal skills are weaker compared to poorer students. In the next section, I illustrate that these two skills are not the only determinants of pricing. So, I do not expect this prediction to hold on an individual basis. Student composition is a direct consequence of the pricing mechanism. Before turning to a formal analysis of the this mechanism, I present some summary evidence regarding distribution of relative ability by income. In Table 2, I present SAT math and verbal averages according to financial aid application status. These averages are almost same for aid applicants and recipients, but the median need of those denied financial aid is significantly higher. This is consistent with the model: Among two students with identical ability, the school tries to admit the one with higher willingness to pay. The means for those who did not apply for aid are given as 731 and 625. The aid applicants have lower math and higher verbal abilities on average, as the model predicts.

I also constructed a relative ability measure for every student by taking the ratio of SAT math to SAT verbal. According to my model, this measure should be positively correlated with income. Unfortunately the income information is not available for those who did not apply for financial aid. So instead of income, I used the tuition paid with the assumption that it is correlated with income. I find the correlation between the relative ability measure and price paid to be 0.18 and significantly different from zero. These findings indicate a relation between relative-ability and income as described in Prediction 3 on an aggregate level.

5.2 Pricing of Abilities:

According to the model, in the presence of close competitors, effective marginal cost of a student is a good approximation to tuition:

$$p_s^* \approx V'(k) + \eta^m (\theta^m - b^m_s) + \eta^v (\theta^v - b^v_s)$$

(12)
I assume tuition is measured with an additive error, and abilities are measured without errors. Then I can rewrite equation (12) as:

$$p_s = \beta_0 + \beta_M SATM_s + \beta_V SATV_s + \varepsilon_s$$  \hspace{1cm} (13)

In the theoretical model, a student is admitted as long as he/she is willing to pay a price at least as high as effective marginal cost. There is no upper bound on price. However, that is not observed in practice. Like many other universities, CMU has a maximum tuition, and about one third of students pay that amount. If the maximum tuition is $\bar{p}$, the net tuition a student pays is $p = \bar{p} - g$, where $g$ is the total grant offered by the school. Also CMU does not offer an aid that exceeds the total cost of education, i.e., $p$ cannot be negative. So $p$ is censored at 0 and $\bar{p}$, and there is a one to one relation between $p$ and $g$ which is censored at 0 and $\bar{p}$ as well. I will use the grant amount $g$ rather than tuition $p$, and focus on the financial aid applicants that decide to enroll.

There are student characteristics other than SAT math and SAT verbal that may affect the financial aid decisions. If there is another ability that schools care about, it will enter the equation in the same way as math and verbal abilities do. Such a variable exists in the data set and it measures non-academic skills of a student. Many schools are concerned about increasing the ratio of females and disadvantaged minorities in their student populations, therefore I control for those too.\footnote{Female and Minority Status differ from the characteristics I consider in the model in two ways: First, for any quality component in the model more is better, however for female and minorities there probably is a threshold which would raise diversity concerns (in the opposite direction) if exceeded. Second, students’ willingness to pay for intellectual diversity may differ from their willingness to pay against under-representation of females and minorities, therefore this should probably be modeled with an additional constraint in the school’s problem. However, it is safe to assume that the thresholds mentioned in the first point are usually not attained. Then explicit modeling of schools’ preferences for diversity would imply tuition discounts for females and minorities.} Finally, the financial aid is need based in principle. So I incorporate a need term to control for variation in income. The equation I estimate is:

$$g_s = \beta_0 + \beta_M SATM_s + \beta_V SATV_s + \beta_{NA} NONACAD_s + \beta_{NEED} NEED_s + \beta_{F} FEMALE_s + \beta_{B} BLACK_s + \beta_{H} HISPANIC_s + \varepsilon_s$$ \hspace{1cm} (14)

The model predicts nonnegative coefficients on all variables, and implies $\beta_V$ to be strictly greater than $\beta_M$.\footnote{Female and Minority Status differ from the characteristics I consider in the model in two ways: First, for any quality component in the model more is better, however for female and minorities there probably is a threshold which would raise diversity concerns (in the opposite direction) if exceeded. Second, students’ willingness to pay for intellectual diversity may differ from their willingness to pay against under-representation of females and minorities, therefore this should probably be modeled with an additional constraint in the school’s problem. However, it is safe to assume that the thresholds mentioned in the first point are usually not attained. Then explicit modeling of schools’ preferences for diversity would imply tuition discounts for females and minorities.
For some applicants, the aid package has a component called “reaction gift.” CMU has a policy that allows applicants to ask for a better financial aid offer. To be eligible for that, the applicant must be accepted at another institution, and must provide a proof of the financial aid offer from that institution. Then for some of those applicants CMU offers a better financial aid package, and reaction gift is that increment. Whether or not the reaction gift is offered depends on the applicant’s and other school’s characteristics, as well as the aid offered from the other school.

The initial offer can be considered as a better indicator of the school’s financial aid policy, since reaction gift awards a signal received after the initial offer. On the other hand, one can interpret the initial offer to be made conservatively, thinking that applicants that deserve more would ask for the reaction offer. In both cases, the role of the reaction offer on enrollment cannot be ignored, so I present results with both the initial offer, and offer with reaction, used as the amount of gift, controlling for asking a reaction in the latter case.

Table 3 summarizes the main findings of my empirical analysis. The dependent variable is total grant amount in Columns A and B. This amount includes the reaction award. Column A gives the OLS estimates. Column B displays the Tobit estimates. The dependent variable in Columns C and D is the initial aid offer before the reaction. Again column C and D provide the OLS and Tobit estimates respectively. Although the coefficients vary a little from column to column, the underlying pattern is relatively robust: Verbal ability is rewarded substantially more than math ability. In fact, the coefficient on SAT math score is insignificant in all four columns. This finding supports my prediction about the pricing of abilities: The rare type of ability in a school is awarded more.

Two other observations help complete the picture of the cross subsidization mechanism. First, the standard deviation of verbal ability (81) is higher than that of math ability (61). This implies larger

---

13 A formal model of this multi-stage process determining financial aid is presented in Epple, Romano, Sarpeca, and Sieg (2006).

14 Verbal ability is not rare in general. Paglin and Rufolo (1990) study the determinants of differences in earnings in the labor market, and provide evidence on the scarcity of quantitative ability (relative to verbal). Their findings suggest that the higher premium on math ability (documented also by James et al. (1989), Grogger and Eide (1995), and Hamermesh and Donald (2006)) is determined by this relative scarcity.

15 The standard deviations are 83 and 67 after correcting for censoring at 800.
variation in awards associated with verbal ability than with math ability through the pricing function (12), even if their prices were same. The fact that verbal ability has a considerably high price strengthens that statement further. Second, low-verbal/high-math profile is associated with higher income students as indicated earlier. This implies the direction of subsidies is from higher income to lower income types, as predicted.

Among other ability related variables, I found nonacademic factor and high school rank to be important components of financial aid. The estimates for nonacademic factor range from 2189 to 2899. I found no significant effect of high school GPA on financial aid. I interpret high school rank as a qualitative variable. I tried different percentiles and found top 10 percent rank to be significant. Almost half of the aid recipients fall in that range, and the estimates range from 1123 to 1528.

Another major component of financial aid is student’s ability to pay. The coefficient of need indicates that, independent of other variables, half of a student’s need is matched on average. When the dependent variable includes the reaction gift, my estimates suggest that applicants who asked for a second offer usually receive a significant increase.

Most schools are concerned about under-representation of minorities and women, so they provide significant tuition discounts to those groups. My estimates indicate that CMU is one of them. I found that African-Americans and Native-Americans receive a discount around ten thousand dollars. The discount is around seven thousand dollars for Hispanics. I did not find a significant discount for Asian or Other-Nonwhite categories. My estimates of the discount for women ranges from 854 to 1325 dollars.

In the theoretical model, I assume that the peer effects are university wide. A particular student’s (negative or positive) contribution affects every other student, and is priced accordingly. However, the distribution of abilities vary significantly among colleges. If instead, peer effects are thought to be college wide, then pricing would be according to contribution to the individual college rather than the entire university. Let \( \text{col}(s) \) be the college in CMU that student \( s \) attends. I transform the ability variables to their deviations from college means as opposed to university means. So an alternative specification is the
following equation:

\[ g_s = \beta^0 + \beta^m(SAT^m_s - \theta^m_{col(s)}) + \beta^v(SAT^v_s - \theta^v_{col(s)}) + \beta^{NA}(NONACAD^m_s - \theta^{NA}_{col(s)}) + \beta^n NEED_s + \epsilon_s \]

for student \( s \), where the overline represents the transformed variable. The Tobit estimates are given in Table 4 with initial offer as dependent variable. The estimates are almost identical to those given in column D of Table 3. Alternative measures such as SAT II subject test scores for math and writing gave similar results, which I do not report here.

In the achievement function, equal weights are assumed for mathematical and verbal abilities. Findings suggest that mathematical ability is a good determinant of future earnings, and verbal ability is not so significant.\(^{16}\) If this is true, students may place more emphasis on peer quality in math. Then rare ability need not necessarily have a higher shadow price in absolute terms. In the case of CMU for example, the prediction tells us that CMU will award verbal ability more (and math ability less, similarly) than a school that specializes in the opposite direction, with higher peer quality in verbal skills and lower quality in math skills. I found that CMU awards verbal ability considerably more than it awards math ability in absolute terms. So if the mentioned argument is valid, it strengthens the findings here.

6 Concluding Remarks

When schools compete for students, specialization may arise as a competition strategy. According to the analysis in this paper, a school that is stronger in one area relative to another tries to attract good students to its weaker area by providing them higher tuition discounts. Among target students, those with lower incomes are more responsive to such an offer, sacrificing fit in exchange for a lower tuition. By contrast, higher income students of that ability type choose to attend a school that is a better fit to their type, and pay higher tuition. In equilibrium, then, the higher income students attend a school that fits their ability

\(^{16}\)See James et al. (1989), Paglin and Rufolo (1990), Grogger and Eide (1995), and Hamermesh and Donald (2006).
type. They subsidize the lower income students who are strong in a different ability type through the pricing (financial aid) mechanism. The latter students then improve the peer group and the quality of education in the secondary specialization of the school, enhancing the educational experience of students in the school’s primary area of specialization. Findings of my empirical analysis support the model’s predictions with respect to the relationship among student abilities, area of specialization in the university, and financial aid awarded. Several recent studies criticize the commonly used peer effects specifications. The multidimensional non-uniform peer effects specification in this paper provides a complementary approach.

Specialization in education often starts much earlier than college. In the US, enrollments in elementary magnet schools and magnet high schools are increasing every year. Many charter schools and preparatory academies also have specialized curricula. Other countries have similar institutions as well (such as specialist schools in the UK). The analysis in this paper can be adapted to address issues related to specialized elementary and secondary education.

In the model developed here, the only contribution of a school to one’s achievement is peer quality. Another potentially important determinant of educational quality is expenditure per student. If expenditure is uniform across students, this paper’s findings remain valid. It is possible, however, that a school’s cost function exhibits returns to specialization. For example, an investment in engineering labs may decrease the cost of further quality increases in that area, by attracting both good students and research funds from government organizations and corporations. If such returns are present, then modifying the cost function used here would imply a stronger structure on quality distribution. I leave these issues for future research.

References:


25


Proof of Proposition 1: Suppose \( \theta^m_i = \theta^m_j \) and \( \theta^v_i = \theta^v_j \) for some \( i \neq j \) s.t. \( i, j \neq 0 \). Condition (7) implies \( p_i(b^m, b^v, y) = p_j(b^m, b^v, y) \) for all \((b^m, b^v, y)\). Condition (8) and market clearance imply

\[
EMC_i (b^m, b^v) = p_i (b^m, b^v, y) \geq p_j (b^m, b^v, y) = EMC_j (b^m, b^v)
\]

for those students who attend school \( j \), and

\[
EMC_j (b^m, b^v) = p_j (b^m, b^v, y) \geq p_i (b^m, b^v, y) = EMC_i (b^m, b^v)
\]

for those who attend school \( i \). Linearity of \( EMC \) implies one of the following three cases: (a) \( EMC_i (b^m, b^v) = EMC_j (b^m, b^v) \) for all \((b^m, b^v)\). (b) \( EMC_i (b^m, b^v) > EMC_j (b^m, b^v) \) for all \((b^m, b^v)\) (c) \( EMC_i (b^m, b^v) \gtrless EMC_j (b^m, b^v) \) for some \( c_1 b^m + c_2 b^v \gtrless \mathcal{C} \), for some constants \( c_1, c_2, \mathcal{C} \). I can rule out (c) since then school qualities would not be same. Also (b) can be ruled out since school \( j \) would be out of market. So the following condition holds:

\[
p_i(b^m, b^v, y) = p_j(b^m, b^v, y) = EMC_i(b^m, b^v) = EMC_j(b^m, b^v) \quad (A1)
\]

Now need to show that types \((b^m_1, b^v_1, y_1)\) and \((b^m_2, b^v_2, y_2)\) exist with \( \alpha_j(b^m_1, b^v_1, y_1) \in (0, 1] \) and \( \alpha_i(b^m_2, b^v_2, y_2) \in (0, 1] \) s.t. school \( i \) can increase profits by admitting the same number of \((b^m_1, b^v_1, y_1)\) types as it expels \((b^m_2, b^v_2, y_2)\) types. I have:

\[
\pi^i_1 = \frac{\partial \pi^i}{\partial [\alpha_i(b^m_1, b^v_1, y_1), f(b^m_1, b^v_1, y_1)]} = p^i_1(b^m_1, b^v_1, y_1, \theta^m_i, \theta^v_i) - V'(k_i)
\]

\[
+ (b^m_1 - \theta^m_1) \frac{1}{k_i} \int_s \int_s \int_s \frac{\partial p^i_1}{\partial \theta^m_i} \alpha_i f db^m db^v dy
\]

\[
+ (b^v_1 - \theta^v_1) \frac{1}{k_i} \int_s \int_s \int_s \frac{\partial p^i_1}{\partial \theta^v_i} \alpha_i f db^m db^v dy
\]
\[(\pi_2^i \text{ is analogous}) \quad \text{and} \quad \pi_{11}^i = \frac{\partial^2 \pi_i}{\partial^2 \alpha_i (b_1^m, b_1^v, y_1) f (b_1^m, b_1^v, y_1)} = -V'' (k_i) + 2 \frac{(b_1^m - \theta_1^m)}{k_i} \left[ \frac{\partial^2 p^*_i (b_1^m, b_1^v, y_1, \theta_1^m, \theta_1^v)}{\partial \theta_1^m} \alpha_i, f \right. \left. db^m db^v dy \right] \]

\[+ 2 \frac{(b_1^v - \theta_1^v)}{k_i} \left[ \frac{\partial^2 p^*_i (b_1^m, b_1^v, y_1, \theta_1^m, \theta_1^v)}{\partial \theta_1^v} \alpha_i, f \right. \left. db^m db^v dy \right] \]

\[+ \left( \frac{(b_1^m - \theta_1^m)}{k_i} \frac{1}{k_i} \right)^2 \int \int \int \frac{\partial^2 p^*_i}{\partial^2 \theta_1^m} \alpha_i, f db^m db^v dy \]

\[+ \left( \frac{(b_1^v - \theta_1^v)}{k_i} \frac{1}{k_i} \right)^2 \int \int \int \frac{\partial^2 p^*_i}{\partial^2 \theta_1^v} \alpha_i, f db^m db^v dy \]

\[+ 2 \frac{(b_1^m - \theta_1^m)}{k_i} \left( \frac{(b_1^v - \theta_1^v)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^m \partial \theta_1^v} \alpha_i, f db^m db^v dy \]

\[(\pi_{22}^i \text{ is analogous}) \quad \text{and} \quad \pi_{12}^i = \frac{\partial \pi_i}{\partial [\alpha_i (b_1^m, b_1^v, y_1) f (b_1^m, b_1^v, y_1)]} = -V'' (k_i)

\[+ \frac{(b_1^m - \theta_1^m)}{k_i} \frac{\partial^2 p^*_i (b_1^m, b_1^v, y, \theta_1^m, \theta_1^v)}{\partial \theta_1^m} + \frac{(b_1^v - \theta_1^v)}{k_i} \frac{\partial^2 p^*_i (b_1^m, b_1^v, y, \theta_1^m, \theta_1^v)}{\partial \theta_1^v}

\[\left. \right. \left. - \left( \frac{(b_1^m + b_1^v - 2 \theta_1^m)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^m} \alpha_i, f db^m db^v dy \right. \]

\[\left. \left. - \left( \frac{(b_1^m + b_1^v - 2 \theta_1^v)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^v} \alpha_i, f db^m db^v dy \right. \]

\[\left. \left. + \frac{(b_1^m - \theta_1^m)}{k_i} \frac{\partial^2 p^*_i (b_1^m, b_2^v, y_1^m, \theta_1^v)}{\partial \theta_1^m} + \frac{(b_1^v - \theta_1^v)}{k_i} \frac{\partial^2 p^*_i (b_1^m, b_2^v, y_1^v, \theta_1^v)}{\partial \theta_1^v} \right. \]

\[\left. \left. + \frac{(b_1^m - \theta_1^m)}{k_i} \left( \frac{(b_1^v - \theta_1^v)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^m} \alpha_i, f db^m db^v dy \right. \]

\[\left. \left. + \frac{(b_1^v - \theta_1^v)}{k_i} \left( \frac{(b_1^v - \theta_1^v)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^v} \alpha_i, f db^m db^v dy \right. \]

\[\left. \left. + \left( \frac{(b_1^m - \theta_1^m)}{k_i} \right) \left( \frac{(b_1^v - \theta_1^v)}{k_i} \right) \int \int \int \frac{\partial^2 p^*_i}{\partial \theta_1^m \partial \theta_1^v} \alpha_i, f db^m db^v dy \right. \]

\[28\]
Let $\Delta_i$ be the number of types $(b_i^m, b_i^v, y_i)$ enrolled in school $i$. For $\Delta_i$ sufficiently small, Taylor’s theorem implies the sign of change in school $i$’s profits $\Delta \pi_i$, will be the same as the sign of:

$$\pi_i \Delta_1 + \pi_2 \Delta_2 + \frac{1}{2} \left[ \pi_{11} (\Delta_1)^2 + \pi_{22} (\Delta_2)^2 + 2 \pi_{12} \Delta_1 \Delta_2 \right]$$

Since prices equal to effective marginal costs $\pi_i^1 = \pi_i^2 = 0$. I will consider admission changes such that $\Delta_1 = -\Delta_2$, so the above equation simplifies to

$$\Delta_i \frac{1}{2} \left[ \pi_{11} + \pi_{22} - 2 \pi_{12} \right]$$

Substituting and simplifying gives $\text{sign}\{\Delta \pi_i\}$

$$= \text{sign}\{ \frac{2}{k_i} (b_i^m - b_i^m) \left[ \frac{\partial p_i^*}{\partial \theta_i^m} (b_i^m, b_i^v, y_i, \theta_i^m, \theta_i^v) - \frac{\partial p_i^*}{\partial \theta_i^m} (b_i^m, b_i^v, y_i, \theta_i^m, \theta_i^v) \right] \}
+ \frac{2}{k_i} (b_i^m - b_i^m)^2 \int \int \int \int \frac{2}{\partial \theta_i^m} \alpha_i f db_i^m db_i^v dy
+ \frac{2}{k_i} (b_i^m - b_i^m) (b_i^m - b_i^m) \int \int \int \int \frac{2}{\partial \theta_i^v} \alpha_i f db_i^m db_i^v dy
+ \frac{2}{k_i^2} (b_i^m - b_i^m) (b_i^m - b_i^m) \int \int \int \int \frac{2}{\partial \theta_i^v} \alpha_i f db_i^m db_i^v dy \}

Suppose that $y_1$ is greater than $y_2$ but very close to it, and that $b_i^m$ is greater than $b_i^m$ but closer still to $b_i^m$, formally, of lower-order difference. Suppose $b_1^m = b_2^m$. Then the above term reduces to:

$$= \text{sign}\{ \frac{2}{k_i} (b_1^m - b_2^m) \left[ \frac{\partial p_i^*}{\partial \theta_i^m} (b_1^m, b_1^v, y_i, \theta_i^m, \theta_i^v) - \frac{\partial p_i^*}{\partial \theta_i^m} (b_2^m, b_2^v, y_i, \theta_i^m, \theta_i^v) \right] \}
+ \frac{1}{k_i^2} (b_2^m - b_2^m)^2 \int \int \int \int \frac{\partial p_i^*}{\partial \theta_i^v} \alpha_i f db_i^m db_i^v dy \}

Now I will show that school $i$ can substitute students this way. From (7),

$$\frac{\partial p_i^*}{\partial \theta_i^m} = \frac{\partial U}{\partial \theta_i^m}$$

Lemma 4 presented below shows that $p_i^* (b_i^m, b_i^v, y, \theta_i^m, \theta_i^v)$ is continuous in $(b_i^m, b_i^v, y)$. By (SCY) and this lemma the first term in the above equation is positive. Moreover it will dominate the second term due to the lower-order of the difference $(b_i^m - b_i^m)$ and again using the same lemma.
Lemma 1: If $\alpha_i(b,y) > 0$, then:

$$U(y - p_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) = U(y - EMC_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v))$$

$$\geq U(y - EMC_k(b^m, b^v), a(\theta_k^m, \theta_k^v, b^m, b^v))$$ for some $j \neq i$ and for all $k \neq i$ s.t. $i, j, k = 1, 2, \ldots, n$.

Proof: Suppose there is a $k \neq i$ s.t.

$$U(y - EMC_j(b^m, b^v), a(\theta_j^m, \theta_j^v, b^m, b^v)) < U(y - EMC_k(b^m, b^v), a(\theta_k^m, \theta_k^v, b^m, b^v))$$

then (3) for $i$ implies $U(y - p_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_k(b^m, b^v), a(\theta_k^m, \theta_k^v, b^m, b^v))$ implying $p_k > EMC_k$ for this type.

By (8), this implies $\alpha_k = 1$. Then, since $\alpha_i > 0$, market clearance is violated. This proves the inequality.

Suppose $U(y - p_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - EMC_j(b^m, b^v), a(\theta_j^m, \theta_j^v, b^m, b^v))$ for all $j \neq i$. Then $i$ can increase $p_i(b^m, b^v, y)$ by retaining type $(b^m, b^v, y)$ as students. This contradicts profit maximization, and proves the equality.

Lemma 2: $U^*(b^m, b^v) = \max_{i \in \{1, 2, \ldots, n\}} U(y - EMC_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v))$ for all $(b^m, b^v, y)$.

Proof: Follows immediately from Lemma 1.

Lemma 3: $U^*(b^m, b^v)$ is continuous in $U(b^m, b^v)$.

Proof: For any positive measure in type space $S$ of students, $U^*(b^m, b^v)$ is the upper envelope of a set of continuous functions in $(b^m, b^v)$ by the above lemma. Hence $U^*(b^m, b^v)$ is piecewise continuous over the type space. It is possible there are jumps at boundary points in $S$ between two schools. Suppose $(b^m, b^v, y)$ is such a point and $U(y - p_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_j(b^m, b^v), a(\theta_j^m, \theta_j^v, b^m, b^v))$.

Since the student attends college $i$, $p_i(b^m, b^v) > EMC_i(b^m, b^v)$, and $p_j(b^m, b^v) < EMC_j(b^m, b^v)$ or $\alpha_j(b^m, b^v, y) = 1$ and market does not clear. Then $U(y - EMC_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_j(b^m, b^v), a(\theta_j^m, \theta_j^v, b^m, b^v))$. Those functions are continuous in $(b^m, b^v, y)$ so according to lemma 1, school $j$ cannot admit any students profitably in the vicinity of $(b^m, b^v, y)$, a contradiction.

Lemma 4: $p_i^*(b^m, b^v, y, \theta_i^m, \theta_i^v)$ is continuous in $(b^m, b^v)$.

Proof: This is implied by lemma 3 and (7).
Proof of Proposition 5:

According to Proposition 1, $\theta_1^m$ cannot equal to $\theta_2^m$ since otherwise symmetry would imply two identical colleges. Let college 1 be the one with higher $\theta^m$. Denote a student’s achievement at College $i$ by $a_i$. This symmetric equilibrium implies:

$$a_1 \geq a_2 \text{ when } b^m \geq b^v$$  \hspace{1cm} (B1)

For a student with relatively higher math (verbal) ability, College 1 (College 2) will be first best, College 2 (College 1) will be second best, and not attending college would be the third best in terms of achievement.

For students on a boundary, tuition in either institution is equal to effective marginal cost of the student in that college by Proposition 2. Then, utility comparisons help locate the income level at which a student with ability $(b^m, b^v)$ is indifferent between any two options:

$$y_{ij}^* (b^m, b^v) = \frac{EMC_i a_i - EMC_j a_j}{a_i - a_j}$$  \hspace{1cm} (B2)

where $i,j \in \{0,1,2\}$, $i \neq j$, and $\forall (b^m, b^v, y)\quad EMC_0(b^m, b^v, y) = 0$.

Because no-college option is free, students take $EMC_0$ as zero when making decisions. Those income thresholds, together with relative ability hyperplane, partition the type space as follows:

$$M^h = \left\{(b^m, b^v, y) : b^m > b^v \text{ and } y \geq \max\{y_{12}^* (b^m, b^v), y_{01}^*(b^m, b^v)\}\right\}$$

$$M^l = \left\{(b^m, b^v, y) : b^m > b^v \text{ and } y_{02}(b^m, b^v) < y < y_{12}^*(b^m, b^v)\right\}$$

$$V^h = \left\{(b^m, b^v, y) : b^m < b^v \text{ and } y \geq \max\{y_{12}^* (b^m, b^v), y_{02}^*(b^m, b^v)\}\right\}$$

$$V^l = \left\{(b^m, b^v, y) : b^m < b^v \text{ and } y_{01}(b^m, b^v) < y < y_{12}^*(b^m, b^v)\right\}$$

$$NC = S \setminus (M^h \cup M^l \cup V^h \cup V^l)$$

Those in $M^h$ and $V^l$ are not constrained by income, and will attend the school that is best for them ($M^h \subset A_1$ and $V^l \subset A_2$). The remaining parts of admission spaces $A_1$ and $A_2$ depend on the relation between shadow prices. The symmetry of the distribution of student ability types implies symmetry in shadow prices, i.e. $\eta_i^m = \eta_i^v$. If the shadow price of the rare ability is higher in a college, then $\theta_1^m > \theta_1^v$ implies:

$$EMC_1 \geq EMC_2 \text{ when } b^m \geq b^v$$  \hspace{1cm} (B3)

This means that higher achievement college is more expensive, indicating potential trade-offs between achievement and tuition (Some algebra shows that the choice-switching threshold $y_{12}^* (b^m, b^v)$ is always positive in this case). For $(b^m, b^v, y)$ in $V^l$ or $M^l$, the second-best college in terms of achievement is the
first best choice given the tuition advantage. Therefore $A_1 = M^h \cup V^l$ and $A_2 = V^h \cup M^l$.

On the other hand, if the shadow price of the rare ability is lower in a college, then $\theta_1^m > \theta_1^r$ implies

$$EMC_1 \preceq EMC_2 \text{ when } b^m \geq b^r$$

(B4)

that higher achievement college is also the less expensive. There are no trade-offs, and every student that attend college will attend his/her first best.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT math</td>
<td>710</td>
<td>63</td>
<td>≤ 500</td>
<td>800</td>
</tr>
<tr>
<td>SAT verbal</td>
<td>654</td>
<td>73</td>
<td>≤ 500</td>
<td>800</td>
</tr>
<tr>
<td>Non-academic</td>
<td>2.62</td>
<td>0.54</td>
<td>0</td>
<td>3.86</td>
</tr>
<tr>
<td>Top 10% HS</td>
<td>0.44</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Need (×1000)</td>
<td>17.14</td>
<td>27.06</td>
<td>&lt; -100</td>
<td>38.77</td>
</tr>
<tr>
<td>Female</td>
<td>0.37</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black/Native American</td>
<td>0.10</td>
<td>0.30</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.08</td>
<td>0.26</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: This table displays the means and standard deviations of SAT scores for groups according to financial aid application status.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>SAT math</th>
<th>SAT verbal</th>
<th>Non-academic Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>1161</td>
<td>717</td>
<td>645</td>
<td>2.54</td>
</tr>
<tr>
<td></td>
<td>(61)</td>
<td>(81)</td>
<td>(0.58)</td>
<td></td>
</tr>
<tr>
<td>F.A. Applicants</td>
<td>796</td>
<td>710</td>
<td>654</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>(61)</td>
<td>(73)</td>
<td>(0.54)</td>
<td></td>
</tr>
<tr>
<td>F.A. Recipients</td>
<td>743</td>
<td>711</td>
<td>655</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>(62)</td>
<td>(73)</td>
<td>(0.54)</td>
<td></td>
</tr>
<tr>
<td>Did not apply to F.A.</td>
<td>345</td>
<td>731</td>
<td>625</td>
<td>2.54</td>
</tr>
<tr>
<td></td>
<td>(57)</td>
<td>(93)</td>
<td>(0.65)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Dependent variable is total grant amount in Columns A and B. This amount includes the reaction award. Column A gives the OLS estimates. Column B displays the Tobit estimates. The dependent variable in Columns C and D is the initial aid offer before the reaction. Again column C and D provide the OLS and Tobit estimates respectively. The sample consists of enrolled financial aid applicants. Standard errors are given in parantheses.

<table>
<thead>
<tr>
<th></th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
<th>(D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-16795</td>
<td>-20720</td>
<td>-14121</td>
<td>-19334</td>
</tr>
<tr>
<td></td>
<td>(3158)</td>
<td>(3536)</td>
<td>(3149)</td>
<td>(3643)</td>
</tr>
<tr>
<td>SAT math</td>
<td>5.55</td>
<td>5.40</td>
<td>3.75</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>(3.58)</td>
<td>(3.98)</td>
<td>(3.56)</td>
<td>(4.09)</td>
</tr>
<tr>
<td>SAT verbal</td>
<td>12.48</td>
<td>14.57</td>
<td>11.61</td>
<td>14.80</td>
</tr>
<tr>
<td></td>
<td>(2.79)</td>
<td>(3.13)</td>
<td>(2.77)</td>
<td>(3.22)</td>
</tr>
<tr>
<td>Non-academic</td>
<td>2691</td>
<td>3066</td>
<td>2303</td>
<td>2696</td>
</tr>
<tr>
<td></td>
<td>(364)</td>
<td>(409)</td>
<td>(362)</td>
<td>(421)</td>
</tr>
<tr>
<td>Need</td>
<td>0.46</td>
<td>0.52</td>
<td>0.47</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Female</td>
<td>1098</td>
<td>1056</td>
<td>1462</td>
<td>1519</td>
</tr>
<tr>
<td></td>
<td>(415)</td>
<td>(464)</td>
<td>(413)</td>
<td>(475)</td>
</tr>
<tr>
<td>Black/NativeAm</td>
<td>10017</td>
<td>10290</td>
<td>9561</td>
<td>9791</td>
</tr>
<tr>
<td></td>
<td>(717)</td>
<td>(790)</td>
<td>(715)</td>
<td>(810)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6952</td>
<td>7270</td>
<td>7075</td>
<td>7485</td>
</tr>
<tr>
<td></td>
<td>(818)</td>
<td>(901)</td>
<td>(814)</td>
<td>(922)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.66</td>
<td>-</td>
<td>0.66</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>805</td>
<td>805</td>
<td>805</td>
<td>805</td>
</tr>
<tr>
<td>N-censored</td>
<td>-</td>
<td>125</td>
<td>-</td>
<td>125</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-</td>
<td>-7195.25</td>
<td>-</td>
<td>-6959.80</td>
</tr>
</tbody>
</table>
Table 4: The three ability-related regressors: SAT math, SAT verbal and Non-academic are their deviations from means of the college attended. The dependent variable is the initial aid offer. Tobit estimates are presented.

(A)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-7048</td>
<td>1165</td>
</tr>
<tr>
<td>SAT math</td>
<td>5.31</td>
<td>4.39</td>
</tr>
<tr>
<td>SAT verbal</td>
<td>14.86</td>
<td>3.19</td>
</tr>
<tr>
<td>Non-academic</td>
<td>2517</td>
<td>413</td>
</tr>
<tr>
<td>Top 10% HS</td>
<td>1544</td>
<td>442</td>
</tr>
<tr>
<td>Need</td>
<td>0.53</td>
<td>0.02</td>
</tr>
<tr>
<td>Female</td>
<td>1341</td>
<td>453</td>
</tr>
<tr>
<td>Black/NativeAm</td>
<td>10044</td>
<td>806</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7846</td>
<td>906</td>
</tr>
<tr>
<td>N</td>
<td>796</td>
<td></td>
</tr>
<tr>
<td>Censored</td>
<td>(53)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: This figure depicts the choice-switching income threshold surface $y_{12}^*$ described in equation (B2) and used in Proposition 5. Together with the equal relative ability plane ($b^m = b^v$), and college attendance surfaces ($y_{01}^*$ and $y_{02}^*$) (none of them shown) they construct the partition in (11).
Figure 2: The admission areas are given at four income levels $y = 48, 56, 72, 108$. Each panel represents an income level, and each point in a square represents a student with ability $(b^m, b^v)$ at the indicated income level. The triangular area (shaded by crosses) at the bottom left corner of every panel shows the ability pairs that choose the no-college option. The white area shows the ability pairs that school 1 (higher $\theta^m$ school) admits. The area shaded by dots indicate the admission area of school 2 (higher $\theta^v$ school).