



Multi-dimensional skills, specialization, and oligopolistic competition in higher education

Sinan Sarpça *

Koç University, Rumelifeneri Yolu, Sariyer 34450, Istanbul, Turkey

ARTICLE INFO

Article history:

Received 29 May 2009

Received in revised form 3 March 2010

Accepted 10 June 2010

Available online 17 June 2010

JEL classification:

I21

D58

Keywords:

Higher education

School competition

Peer effects

Specialization

Multi-dimensional skills

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 identical schools compete for students, specialization may arise as a competition strategy. This serves rich students' education goals well. Poorer students, however, may attend schools with specializations that do not cater to their relative strengths. 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.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Recent empirical research in labor economics provides plenty of evidence that skills are multi-dimensional.¹ Theoretical research on peer effects in education has not utilized these findings yet. A common assumption in the related literature is that the quality of a peer group can be captured by a single-dimensional index.² The models that adopt this assumption also cannot 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.

We develop a new model in which students that differ in income and various skills decide between education alternatives, given their preferences over skill composition of their prospective peers and costs of these alternatives. The general equilibrium model incorporates multi-dimensional skills of individuals and specialization of colleges

into a school competition framework, neither of which are considered in previous research. We study schools' optimal admission and pricing policies, students' decisions, and the equilibrium consequences to different types of students. Our approach departs from the related literature also by capturing heterogeneity in students' preferences over schools, and differential structure of peer effects, as opposed to a uniform structure that is commonly adopted in index and linear-in-means models.³

Researchers have been using general equilibrium models to study consequences of competition in primary and secondary education and the likely effects of policy changes (Benabou, 1996, 2002; Caucutt, 2002; deBartolome, 1990; Epple and Romano, 1998; Fernandez and Rogerson, 1998, 2003; Nechyba, 2000). The single-dimensional school quality assumption in these papers is the natural starting point, and has proved very useful for studying many questions. However, there are other interesting topics (multiple skills, specialization, etc.) for which this assumption is not suitable. It implies a strong substitutability relation between many different attributes a school provides in reality. In these models, the competition for students results in a strict hierarchy of schools measured by school quality, which is common to all students. There is never a mismatch between a student's skills and

* Tel.: +90 212 338 1214.

E-mail address: ssarpca@ku.edu.tr.

¹ The idea dates back to Roy (1951). See, for example, Acemoglu (2002), Bowles et al. (2001), Cunha and Heckman (2007), and Cunha et al. (2006).

² See Benabou (1996), Caucutt (2002), deBartolome (1990), Epple and Romano (1998, 2003), and Nechyba (2000).

³ Index models assume that all effects of a student's peers on the student's own outcomes can be summarized by a scalar. The linear-in-means models assume that a student's outcome is a linear function of the mean of his/her peers' outcomes.

a school's attributes in equilibrium, and all students in a school benefit from the peer quality in the school in a uniform or linear way. [Epple et al. \(2003, 2006a\)](#) also construct a formal model of the market for higher education and study the quality hierarchy in the market. This paper tackles a completely different question—multi-dimensional skills and specialization—and the model in this paper differs from theirs and other models cited above in several important ways. First, we allow school quality to be multi-dimensional. This allows schools to compete by horizontal differentiation, or specialization. Second, we allow students to have multiple skills. This raises a concern for “fit” to a school's profile, and allows for heterogeneity in student preferences over schools. The model in this paper also provides an explanation to why some students choose schools that are poorly suited to their relative talents. Third, we introduce multi-dimensional peer effects to deal with multiple skills of students. These multi-dimensional peer effects operate non-uniformly, through interaction with an individual's own characteristics. Peer effects in education have been discussed since the Coleman Report (1966), however the empirical evidence remains controversial to date.⁴ Several recent studies criticize the commonly used specifications.⁵ The specification in this paper provides an alternative approach.

The formal model suggests that specialization of schools can arise in equilibrium, and yields predictions about financial aid policies and student compositions of specializing colleges.⁶ We 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. We allow students to benefit both from having peers with abilities and interests that are congruent with their own and the complementing effects of having a peer group with different strengths. Specialization makes a college more appealing to students who have relatively strong skills in the area in which it specializes, however, students and schools may prefer to avoid the pitfall of an overly narrow education. If students demand complementing peers as well as specialization, this creates a challenge for the school in attracting able students for whom there are other schools with specializations that fit them better. Financial aid is a tool that a college may attempt to use to attract such students. The model predicts that schools' financial aid policies will differentiate between different types of skills, offering more financial aid for skills that are harder to attract via the school's specialization. Lower-income students are more responsive to such offers. Our model predicts a cross-subsidization mechanism that is very different than other market models: Able high-income students attend colleges with specializations congruent with their interests and receive little financial aid. By contrast, able low-income students 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.

Since equilibria of this model can only be computed numerically, we use a quantitative specification of our model to gain further insight into the properties of equilibrium. Our results suggest that when a school's financial aid policy favors a particular type of skill, this will attract students that are exceptionally strong in that type of skill, establishing a relation between relative abundance of a type of skill

and its variation in the school. The theory does not rule out a specialization equilibrium in which students choose schools according to their relative strengths only. Income variation in the population is a key variable in this, and computational analysis suggests that only a modest income variation generates the equilibrium and cross-subsidization described in the previous paragraph, with students choosing schools based on their incomes as well as relative strengths.

Finally, we provide an empirical analysis that complements the theoretical analysis. Using a new data set on the entire set of students at a specializing university, we investigate if the school's financial aid policy differentiates between different types of skills, whether a student's fit to school profile changes with income, and whether the within-school variation of different skills change according to the school's specialization. The results of this analysis provide strong support for the financial aid predictions of the model, and suggest 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.

The rest of the paper is organized as follows: [Sections 2 and 3](#) describe the theoretical framework. The data are discussed in [Section 4](#). [Section 5](#) provides a number of predictions about equilibrium implications and 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 $\mathbf{b} = (b^m, b^v)$.⁷ 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 $S \subseteq (0, b_{max}]^2 \times (0, y_{max}]$.

An applicant has the utility function $U(\cdot)$ increasing in its two components: numeraire consumption and educational achievement. Consumption is equal to the income after tuition expenditure p . Educational achievement is increasing in student's own capabilities $\mathbf{b} = (b^m, b^v)$ as well as the attended school's quality. There are a finite number of competitive schools ($i = 1, \dots, n$). A school's quality is determined by the quality of the peer group $\theta_i = (\theta_i^m, \theta_i^v)$, where each component represents the mean level of that type of ability in the school. So the utility of an applicant is given by $U(y - p, a(\theta_i^m, \theta_i^v, b^m, b^v))$ 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 $\theta_0 = (\theta_0^m, \theta_0^v)$.⁸ We assume θ_0^m equals θ_0^v and college qualities exceed those of the outside alternative, i.e., $\min_{i, j \in \{1, \dots, n\}} \{\theta_i^m, \theta_j^v\} \geq \theta_0^m = \theta_0^v$.⁹

⁴ [Manski \(1993\)](#) discusses the problems in identification. Recent empirical work on peer effects in higher education includes [Zimmerman \(2003\)](#), [Sacerdote \(2001\)](#), [Stinebrickner and Stinebrickner \(2006\)](#), [Winston and Zimmerman \(2004\)](#), and [Arcidiacono and Nicholson \(2005\)](#).

⁵ [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 model.

⁶ The predictions this paper derives and tests 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, the market implications will be identical.

⁷ The vector b is two dimensional mainly for expositional convenience. First four propositions can immediately be generalized for n -dimensional vectors.

⁸ 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.

⁹ This is for simplicity of exposition. Since we consider only desirable peer characteristics here, the inequality could be obtained as an equilibrium property even if it was not assumed.

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

$$\partial \left(\frac{\partial U / \partial a}{\partial U / \partial y} \right) / \partial y > 0 \tag{SCy}$$

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:

$$C(k) = V(k) + F \tag{1}$$

where k denotes the size of the college, $V' > 0$ and $V'' > 0$, and k^* denotes the “efficient scale:” $k^* = \arg \min_k C(k) / k$. These assumptions assure that, in equilibrium, there will be a finite number of schools.

The colleges maximize profits by choosing tuition and admission policies behaving as utility takers.¹⁰ 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. The equilibrium number of schools is the number that can profitably enter. 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:

$$\max_{\theta_i^m, \theta_i^v, k_i, \alpha_i(b^m, b^v, y), p_i(b^m, b^v, y)} \int_S \int_S p_i(b^m, b^v, y) \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy - V(k_i) - F \tag{2}$$

subject to the constraints:

$$U(y - p_i(b^m, b^v, y), a(\theta_i^m, \theta_i^v, b^m, b^v)) \geq \max_{j \in \{0, 1, \dots, n \mid j \neq i; \alpha_j(b^m, b^v, y) > 0 \text{ is in the optimal set of } j\}} U(y - p_j(b^m, b^v, y), a(\theta_j^m, \theta_j^v, b^m, b^v)) \quad \forall (b^m, b^v, y); \tag{3}$$

$$\alpha_i(b^m, b^v, y) \in [0, 1] \quad \forall (b^m, b^v, y); \tag{4}$$

$$k_i = \int_S \int_S \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy; \tag{5}$$

$$\theta_i^m = \frac{1}{k_i} \int_S \int_S b^m \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy; \tag{6}$$

$$\theta_i^v = \frac{1}{k_i} \int_S \int_S b^v \alpha_i(b^m, b^v, y) f(b^m, b^v, y) db^m db^v dy;$$

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) determine the college's size and quality vector.

Definition. An equilibrium consists of a set of admission and pricing functions $\alpha_i(b^m, b^v, y)$ and $p_i(b^m, b^v, y)$ for each college; a vector of college characteristics $(\theta_i^m, \theta_i^v, k_i)$ for each college; utility levels $U^*(b^m,$

$b^v, y)$ for each student type; and an allocation of students into the n colleges and no-college such that:

- (i) Every student (b^m, b^v, y) is allocated to a preferred option in his/her effective choice set (the colleges that admit the student in equilibrium and the no-college option), taking colleges' pricing/admission functions and qualities as given:

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

- (ii) Each college chooses its size, peer quality vector, admission and tuition policies, to maximize profits, taking as given other colleges' admission and pricing policies, and the resulting maximum alternative utility levels for students:

$$[\theta_i^m, \theta_i^v, k_i, p_i(b^m, b^v, y), \alpha_i(b^m, b^v, y)] \text{ satisfy (2)–(6) for } i = 1, 2, \dots, n. \tag{PM}$$

- (iii) Schools can freely enter and exit.
- (iv) Every student will either attend a college or be matched to the no-college option, i.e., markets clear:

$$\sum_{i=0}^n \alpha_i(b^m, b^v, y) = 1 \quad \forall (b^m, b^v, y). \tag{MC}$$

3. Theoretical results

3.1. Properties of equilibrium

The first-order conditions for the school's problem described in Eqs. (2)–(6) are as follows:

$$U(y - p_i^*, a(\theta_i^m, \theta_i^v, b^m, b^v)) = U^*(b^m, b^v, y) \tag{7}$$

$$\alpha_i \begin{cases} = 0 \\ \in [0, 1] \\ = 1 \end{cases} \text{ as } p_i^*(b^m, b^v, y, \theta_i^m, \theta_i^v) \begin{cases} < \\ = \\ > \end{cases} \underbrace{V(k_i) + \gamma_i^m(\theta_i^m - b^m) + \gamma_i^v(\theta_i^v - b^v)}_{EMC_i(b^m, b^v)} \tag{8}$$

$$\gamma_i^m = \frac{1}{k_i} \int_S \int_S \frac{\partial p_i^*}{\partial \theta_i^m} \alpha_i f db^m db^v dy \tag{9}$$

$$\gamma_i^v = \frac{1}{k_i} \int_S \int_S \frac{\partial p_i^*}{\partial \theta_i^v} \alpha_i f db^m db^v dy.$$

Condition (7) describes a school's optimal tuition policy. The RHS term is the student's reservation utility as described by 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 Eq. (9) calculate these marginal effects deriving from changes in θ_i^m and θ_i^v . Consider the RHS term in Eq. (8). The first component is the marginal cost that comes from cost function in Eq. (1). The second and third terms represent the marginal cost arising from the student's (negative or positive) contribution to peer quality. For example if a student's ability b^m is below the mean ability θ^m , the student pays $\gamma_i^m \times (\theta_i^m - b^m)$ to compensate for the negative contribution, whereas if its above θ^m , the student receives the same amount in tuition discounts for

¹⁰ See Scotchmer (1994). In Epple and Romano (1998) schools maximize profits also. An alternative objective used in the literature is quality maximization (Epple et al. (2003, 2006a)). We discuss below that a single dimensional quality measure for schools is not appropriate for this setup.

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.¹¹

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 $i, j \in \{0, 1, \dots, n\}$ s.t. $i \neq j$).*

The proof is in Appendix A. 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.¹² 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 horizontally, by specialization.

Let $A_i \equiv \{(b^m, b^v, y) \in S | \alpha_i(b^m, b^v, y) > 0\}$ is optimal when $p_i(b^m, b^v, y) = EMC_i(b^m, b^v)$ 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, and 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) *The equilibrium allocation of students into schools is as if schools set tuitions equal to equilibrium effective marginal costs 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 Eq. (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 $a(\theta_i^m, \theta_i^v, b^m, b^v) \leq a(\theta_j^m, \theta_j^v, b^m, b^v)$ 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. Adopting a single-dimensional quality measure in this framework simplifies the analysis and strengthens the results, but proves to be overly restrictive. Consider a quality function $q_i = q(\theta_i^m, \theta_i^v)$ (e.g., a weighted average), where achievement of a student depends on a scalar q but not on individual θ 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: $q_n > q_{n-1} > \dots > q_1$.*

¹¹ The first-order conditions are sufficient for a maximum when $\alpha_i(b^m, b^v, y)$ is 0 or 1. When $\alpha_i(b^m, b^v, y) \in (0, 1)$, a cost function with a large curvature (high V'') is sufficient for a local maximum. A proof is available upon request from the author.

¹² The cost function (1) ensures the equilibrium number of schools is finite, so there is always room for substantial differentiation.

This is a direct corollary of Proposition 1 under the additional structure. It shows that a one-dimensional framework (e.g. Epple and Romano, 1998) is not good for studying specialization (horizontal differentiation), a multi-dimensional quality measure must be retained. The next section presents one such specification and discusses some alternatives.

3.2. Specialization

The main goal of this paper is to study the implications of specialization. Whether specialization can be observed in equilibrium depends on the specification of the achievement function. For instance, if school quality enters the achievement function as a single-dimensional variable $a(\theta^m, \theta^v; b^m, b^v) = a(q; b^m, b^v)$ for $q \in \mathbb{R}_+$, then, for a given quality level, students are indifferent to its composition, since different components are either easily substitutable (e.g., $q = (\theta^m)^\beta (\theta^v)^\gamma$, $q = \beta\theta^m + \gamma\theta^v$, for $\gamma, \beta > 0$) or not substitutable at all (e.g., $q = \max\{\beta\theta^m, \gamma\theta^v\}$). This class of specifications implies a strict hierarchy of school qualities, and all students sharing the same preferences over schools. These implications are inconsistent with observed facts such as different types of schools, and variation in students' preferences over these schools. Analysis of specialization requires that we allow for richer interaction between a student's own abilities and peer abilities. Hence, we propose and use the following achievement function, with many plausible implications that we discuss below:

$$a(\theta^m, \theta^v; b^m, b^v) = [(\theta^m)^\beta (b^m)^\gamma + (\theta^v)^\beta (b^v)^\gamma]^\frac{1}{\delta} \quad (10)$$

$$0 \leq \beta \leq \gamma \leq 1; \gamma \leq \delta.$$

The bounds on the parameter values imply that this function is increasing in its four components. The choice of $\gamma \leq \delta$ ensures diminishing returns on all inputs, and $\beta \leq \gamma$ allows own ability to have a greater weight than peer quality on a student's achievement.

Under this specification, a student values a school's overall quality, but also differentiates between its components. The student's preferences over composition of quality depend on the student's own skill distribution. For example, when $\theta^m = \theta^v$, a student with relatively stronger mathematical skills ($b^m > b^v$) prefers an increase in the mathematical quality to an equal increase in the verbal quality. Moreover, among students with the same overall level of achievement, an increase in mathematical quality affects students' achievements in proportion to their own mathematical skill levels. As a result of such properties, students' rankings of colleges exhibit variation, which is determined by the interaction, or fit, between student and college characteristics. For instance, given two schools such that $(\theta_1^m, \theta_1^v) = (\theta', \theta)$ and $(\theta_2^m, \theta_2^v) = (\theta, \theta')$ for $\theta' > \theta$, students with relatively stronger mathematical skills ($b^m > b^v$) benefit more from the first school, whereas students with relatively stronger verbal skills ($b^m < b^v$) benefit more from the second school. This may give schools incentives to differentiate themselves horizontally.

Despite all these, this achievement specification does not prevent substitution between different components of quality. Above we mentioned that when $\theta^m = \theta^v$, a student with relatively stronger mathematical skills ($b^m > b^v$) prefers an increase in θ^m to an equal increase in θ^v . However, the student is not indifferent to changes in θ^v , and a larger increase in θ^v can be preferred over a smaller increase in θ^m . As a result, this student may even prefer a school with verbal specialization over a school with math specialization, if the overall quality is sufficiently higher. Also, the returns to specialization are diminishing. If we increase θ^m keeping θ^v constant, the marginal achievement from θ^m decreases, and the student eventually prefers an increase in θ^v to an equal further increase in θ^m .

A noteworthy feature of this specification is that schools are identical in their cost functions and production functions. The weight

on quality β is same across schools and across different dimensions of quality within each school. Similarly, the weight on student ability γ is same across schools and across different skill types within each school.

These features of the model make 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.¹³

To focus on the effects of competition by specialization, we restrict our attention to the case where only two colleges can cover their costs in equilibrium ($i=1,2$). There is a 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 $U = (y - p)a(\cdot)$. We consider an allocation that satisfies the following five properties: a) market clearing; b) utility maximization; c) consistency of reservation utilities¹⁴; 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 et al. (2006a).¹⁵ 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., $\theta_1^m = \theta_2^v$, and $\theta_1^v = \theta_2^m$.

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 $\Theta_i = (\theta_i^m, \theta_i^v)$ and $EMC_i(b^m, b^v)$ for options $i = 0, 1, 2$, these thresholds $\{y_{ij}^*(b^m, b^v) : ij \in \{0, 1, 2\}, i \neq j\}$ can be mapped by an investigation of boundary students using Proposition 2 (see Appendix A for derivations). Consider the following partition of the student type space S with respect to relative ability and income thresholds described above:

$$\begin{aligned}
 M^h &= \{(b^m, b^v, y) : b^m > b^v \text{ and } y > \max\{y_{12}^*(b^m, b^v), y_{01}^*(b^m, b^v)\}\} \\
 M^l &= \{(b^m, b^v, y) : b^m > b^v \text{ and } y_{02}^*(b^m, b^v) < y < y_{12}^*(b^m, b^v)\} \\
 V^h &= \{(b^m, b^v, y) : b^m < b^v \text{ and } y > \max\{y_{12}^*(b^m, b^v), y_{02}^*(b^m, b^v)\}\} \\
 V^l &= \{(b^m, b^v, y) : b^m < b^v \text{ and } y_{01}^*(b^m, b^v) < y < y_{12}^*(b^m, b^v)\} \\
 NC &= S \setminus (M^h \cup M^l \cup V^h \cup V^l)
 \end{aligned}
 \tag{11}$$

We get a specialization equilibrium where one college is more desirable for half of the population, the other college is more desirable for the other half of the population, 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 ($\eta_i^m = \eta_j^v ; i, j \in \{1, 2\}, i \neq j$). 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 $\{(\theta_1^m, \theta_1^v), (\theta_2^m, \theta_2^v)\}$ are such that $\theta_1^m = \theta_2^v > \theta_1^v = \theta_2^m$. Depending on the model parameters shadow prices can be related in two ways: (i) If $\eta_1^m = \eta_2^v < \eta_2^m = \eta_1^v$, 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 $\eta_1^m =$

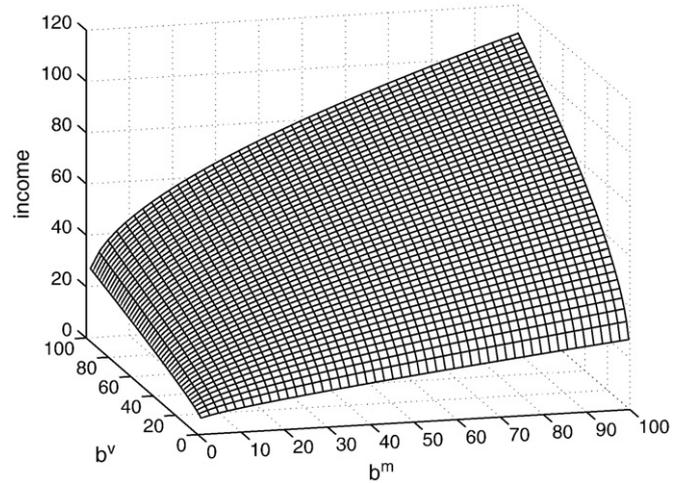


Fig. 1. This figure depicts the choice-switching income threshold surface y_{12}^* described in Eq. (B2) and used in Proposition 5. Together with the equal relative ability plane ($b^m = b^v$), and college attendance surfaces (y_{01}^* and y_{02}^*) (none of them shown) they construct the partition in Eq. (11).

$\eta_1^v \geq \eta_2^m = \eta_2^v$, 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 Appendix A. Proposition 5 implies that students with higher incomes, M^h and V^h , attend College 1 (higher θ^m) and College 2 (higher θ^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 consists 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 ($\eta_1^m < \eta_2^m$ and $\eta_1^v > \eta_2^v$), but can get higher educational achievement at College 1 ($\theta_1^m > \theta_2^m$ and $\theta_1^v < \theta_2^v$). So student's decision depends on whether income is "high enough" to substitute achievement gain for loss from higher tuition. This choice-switching income threshold y_{12}^* is increasing in b^m : An increase in b^m increases the achievement at College 1. However, it also increases the student's achievement at College 2, since the math quality of College 2 is not zero, it is just lower than that in College 1. But note that θ^v is higher at College 2, and also one's own ability matters more for achievement than peer quality ($\beta \leq \gamma$), so the tuition advantage dominates the achievement loss from switching.¹⁶ The threshold y_{12}^* is increasing in b^v too, since an increase in b^v converges the student to $b^m = b^v$, decreasing the achievement gap between schools (This is discussed more below). Fig. 1 illustrates the typical shape of this y_{12}^* surface obtained using a computational model.

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 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

¹³ A specification such as $a(\cdot) = \max\{(\theta^m)^\beta (b^m)^\gamma, (\theta^v)^\beta (b^v)^\gamma\}$ can also generate heterogeneous preferences and specializing schools in equilibrium, without any of the substitutability relations discussed above. The associated equilibrium is qualitatively same as the second type of specialization equilibrium discussed below, so we skip a separate discussion. A specification such as $a(\cdot) = (\theta^m)^\beta (b^m)^\gamma (\theta^v)^\beta (b^v)^\gamma$ is equivalent to a single dimensional quality specification with $q = (\theta^m \theta^v)^\beta$.

¹⁴ This refers to the property that every school takes the utility defined on the RHS of (UM) as the reservation utility in equilibrium.

¹⁵ As explained in Epple et al. (2006a), 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.

¹⁶ The estimates in the literature for β/γ range between 0 and 0.4, and we adopt values in this range.

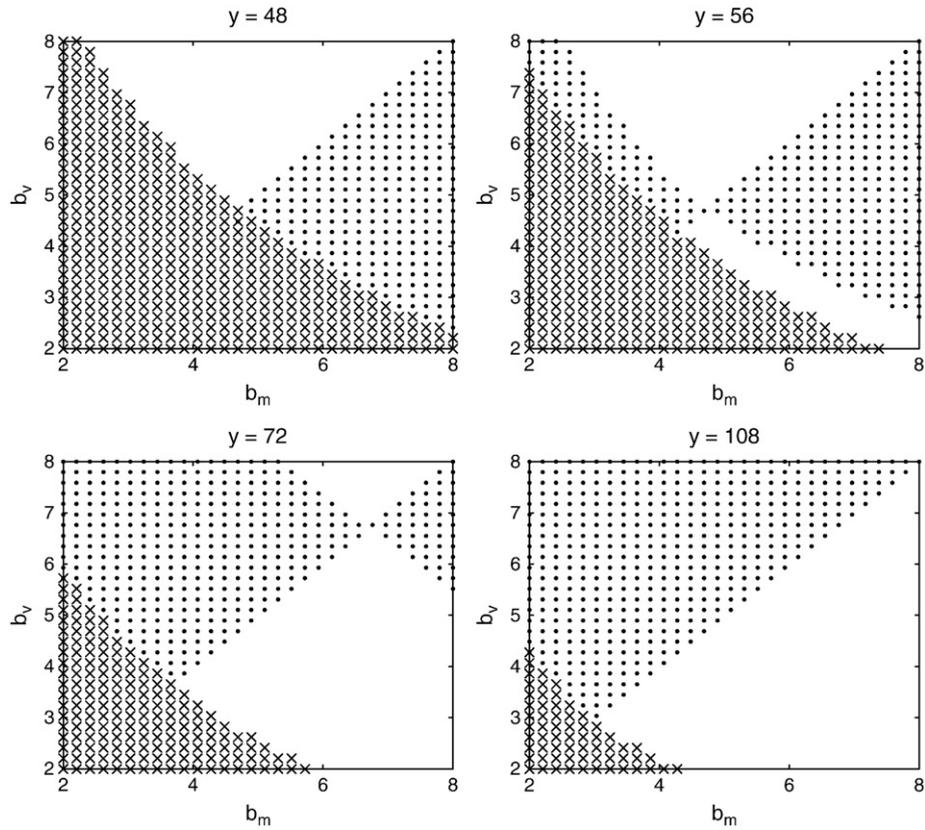


Fig. 2. The admission areas are given at four income levels $y = 48, 58, 72, 108$. Each panel represents an income level, and each point in a square represents a student with ability (b^m, b^v) at the indicated income level. The triangular area (shaded by crosses) at the bottom left corner of every panel shows the ability pairs that choose the no-college option. The white area shows the ability pairs that school 1 (higher θ^m school) admits. The area shaded by dots indicate the admission area of school 2 (higher θ^v school).

the opposite school although they have high incomes (for example, see NE corner in Panel 3 ($y = 72$) of Fig. 2). There are two reasons for that. First, their b^m and b^v are very high compared to θ^m and θ^v at either school, therefore the share of peer quality in achievement is very low. 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 every student attends the college that offers maximum achievement for the student's type, so each school consists of students with similar specialization preferences. Even though it is not possible to establish analytically, computational models suggest that this is a degenerate equilibrium.¹⁷

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 Eq. (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 Fig. 2). This concludes the discussion of Proposition 5.

¹⁷ This equilibrium can be observed, for example, if every household has the same income. We have solved a large number of computational models with different sets of parameters that are consistent with previous studies and observed facts. None of these models generated this type of equilibrium.

We 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. We report one such model below.

3.3. An example

Both ability types are distributed normal with a mean of 510 and a standard deviation of 112 points, censored at 200 at the bottom and at 800 at the top.¹⁸ There is a 0.70 correlation between b^m and b^v , and a 0.25 correlation between income and each ability.¹⁹ The U.S. income distribution is typically approximated using a lognormal specification.²⁰ The mean and median household incomes are \$57,852 and \$42,409 according to 2002 census data. Measured in thousands of dollars, these imply a mean 3.75 and a standard deviation 0.79 for the lognormal distribution.

There are two schools with the same quadratic cost function $C(k) = 150 + 14,000k + 1,389k^2$, and an outside-college option at no cost with quality $\theta_o = \theta_o^m = \theta_o^v = 0.05$, so that: a) One third of the population attends college in equilibrium, given the preference parameters

¹⁸ The SAT was initially scaled to make 500 the mean score on each section with a standard deviation of 100, and was readjusted to this scale in 1995. However its distribution differs slightly from one year to another. College Board reports the means and standard deviations of SAT math and verbal tests as 514 (113) and 506 (111) for 2001. We ignore the few points difference and use the averages to be consistent with the symmetric distributions in the model.

¹⁹ These are approximately the figures reported in Koenig et al. (2008) and Blenden et al. (2007) as correlation between SAT math and verbal scores, and correlations between income and math/verbal skills.

²⁰ See Dragulescu and Yakovenko (2001), Battistin et al. (2009) among many others.

below²¹; b) Average cost is about \$14,400²²; c) Equilibrium number of schools is two; d) Schools operate at efficient scale. The average tuition and fees amount to about a quarter of the average income in the data, so we set $\delta = 3$ to match the budget share of tuition. The estimates in the literature suggest the range of 0.00–0.35 for β/γ .²³ We adopt the value $\beta/\gamma = 0.2$ and set $\beta = 0.12$ and $\gamma = 0.6$. The qualitative results of this analysis are robust to perturbing β , γ , and δ by $\pm 33\%$. Equilibrium values are given in the table below. Admission spaces are depicted in Fig. 2.

Example 1: Computed equilibrium							
	θ^m	θ^v	k	η^m	η^v	σ^m	σ^v
School 1	560	535	0.17	4.20	4.40	99	111
School 2	535	560	0.17	4.40	4.20	111	99

School 1 has a higher type-m quality, and provides higher achievement than school 2 to those with $b^m > b^v$. In school 1, one point of b^m is worth \$4.20 and one point of b^v is worth \$4.40, i.e., rare ability has a higher shadow price. The reason we do not see larger spreads between θ^m and θ^v and between η^m and η^v is partly due to school sizes in this example, and partly implied by the model: As θ^m increases in a school, b^v becomes rare and η^v increases. This attracts high b^v students, and an increase in θ^v will follow. However, the differences between θ^m and θ^v and between η^m and η^v increase as school sizes decrease, 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. At a given school, high-ability students with low incomes are subsidized by low ability–high-income types via tuition discounts due to their contribution to average ability. The model in this paper also exhibits such subsidization across ability levels within each skill type. In addition, the model generates a second form of subsidization across skill types: The students that benefit more from a school's specialization have higher incomes, and on average pay higher tuition compared to other students in the same school who have different relative strengths, and thus could benefit more from attending schools with different specializations. To see this, consider two students $(b_1^m, b_1^v, y_1) = (b + \Delta b, b, y_1)$ and $(b_2^m, b_2^v, y_2) = (b, b + \Delta b, y_2)$ for $\Delta b > 0$, attending school 1 of Proposition 5 ($\theta^m > \theta^v$). In terms of educational achievement, school 1 is first choice for student 1, whereas student 2 prefers the peer composition of school 2 over that of school 1. Even though these two students have the same overall ability, the school charges student 1 a higher tuition than it does to student 2, and $y_1 > y_2$ according to Proposition 5.

Note that the above comparison is between two students with the same overall ability levels, and it does not imply that any student that benefits more from a school's specialization will pay a higher tuition than any fellow student for whom there is a more suitable alternative school. The overall ability levels are also important. Consider again two students attending school 1 of Proposition 5: $(b_1^m, b_1^v, y_1) = (b + \Delta b, b, y_1)$ and $(b_3^m, b_3^v, y_3) = (b', b' + \Delta b, y_3)$ for $b \neq b'$. In terms of educational achievement, school 1 is first choice for student 1, whereas student 3 prefers the peer composition of school 2 over that of

school 1. The school will charge student 1 a lower tuition than it does to student 3 if the first student is sufficiently better overall (if $b > b' + \frac{(\eta^m - \eta^v)}{(\eta^m + \eta^v)} \Delta b$ to be precise). However, student 1 will still pay a higher tuition than his counterpart student 2 as described above.

Our goal in the empirical analysis is to document this cross-subsidization at a specializing school, in the light of the within-school predictions of the model we discuss in Section 5.

4. Data

To test the predictions of the model, we use data on the entire set of applicants at Carnegie Mellon University (CMU hereafter) in the year 2002. Carnegie Mellon is well-suited for our purposes. It is a technical university by reputation. The relative abundance of math ability is evident in the SAT scores. It is highly selective, 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, demographic characteristics, and other information obtained during the admission process. The main file we use contains 8983 applications, 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. Four departments do not require the standardized ability measures for admission, so we exclude those from our analysis.²⁴ Then we have 796 students that do not pay full tuition, that are priced based on their observable characteristics. Table 1 displays the descriptive statistics for our sample.

When describing the model, we 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 the university's standpoint. That allows us 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. We 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 of 645 and standard deviation of 81. As alternative measures, the subject tests SAT II writing test, and SAT II 1C and 2C (math) test scores are available as well.

Income related information is also crucial for our analysis. The data set contains various income and need measures for those applied to financial aid. "Need" is defined as the difference between the student's ability to pay and the total annual cost of education. We 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*. We discuss this in greater detail in the next section. As aid variables we 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 –

²¹ This is approximately the ratio attending four-year degree programs in the U.S. (Epple et al., 2003).

²² NCES reports the average tuition and fees at all four-year institutions as \$14,439 for 2002–2003.

²³ See Henderson et al. (1978), Epple and Romano (1998), and Hoxby and Weingarth (2005) among others.

²⁴ The four departments are Music, Art, Drama, and Design. The admission decisions are based heavily on very specific artistic skills that are evaluated by auditions and portfolio reviews.

Table 1
Summary statistics—enrolled financial aid applicants (N = 796).

	Mean	Std. dev.	Min	Max
SAT math	710	63	≤500	800
SAT verbal	654	73	≤500	800
Non-academic	2.62	0.54	0	3.86
Top 10% HS	0.44	0.50	0	1
Need (× 1000)	17.14	27.06	<−100	38.77
Female	0.37	0.48	0	1
Black/Native American	0.10	0.30	0	1
Hispanic	0.08	0.26	0	1

\$100,000) up to the full cost. Average need is \$9647 and standard deviation is \$37,070.

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

5. Empirical analysis

In this section we focus on within-school implications of specialization. Predictions for comparing schools that differ in their direction of specialization also can be derived from the model, however: 1. We have access to data from just one university; 2. Comparative statements may not always hold if schools that are compared differ in other characteristics except for their specialization, as it is defined in the model.²⁵

We are interested particularly in documenting cross-subsidization discussed in Section 3.4. Using the data described above, we can examine if the school's financial aid policy differentiates between math and verbal scores, whether there is a relation between one's income and fit to school's specialization, whether the within-school variation of math and verbal ability differ, and compare all these to the model's predictions.

A school is more appealing to students who have relatively strong skills in the area in which it specializes. However, to avoid enrolling a particular type of student exclusively, the school may use financial aid policies to attract students with different strengths. The model suggests that a technical school like Carnegie Mellon will offer a higher tuition discount for the verbal ability that is relatively rare (and harder to attract via its specialization) in the school:

Prediction 1. *If $\theta^m > \theta^v$, then $\eta^m < \eta^v$.*

This financial aid policy that favors verbal ability gives the largest financial advantage to prospective students that have large positive deviations in (relative) verbal ability. This may cause the variation in verbal ability to exceed that of the math ability in the school. As noted earlier, this relationship is a robust feature of computed equilibria:

Prediction 2. *If $\theta^m > \theta^v$, then $\sigma^m < \sigma^v$.*

Among students with relatively high verbal ability, those with lower incomes would be more responsive to such a financial aid offer, sacrificing some additional utility from education they could obtain at another type of school in exchange for the low tuition here. Similarly, students with relatively high math ability and low incomes may

²⁵ The model assumes schools are identical in every aspect other than their students, and focuses on peer quality as the single input to school quality. However, schools may also specialize through allocation of expenditures across departments. We comment more on this in the conclusion section.

choose to attend schools with financial aid policies that favor math ability instead. The implication of resulting admission spaces—such as the one in Fig. 2—can be summarized as:

Prediction 3. *The fit between a student's ability profile and school's specialization increases in student's income.*

These three predictions characterize cross-subsidization. The first indicates the financial advantage of attending a school less desirable in terms of its specialization. The second emphasizes the rare ability has a greater role in creating this advantage. The third completes the picture by indicating the direction of subsidies as from rich to the poor.

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—on average—compared to poorer students. As we discuss further in the next section, CMU also places weight on other student attributes such as high school leadership roles and extra-curricular activities, and CMU also values gender and racial diversity. Hence, we do not expect the pricing predictions with respect to verbal and quantitative skills to hold exactly. We do, however, expect that, controlling for other factors, student verbal and quantitative skills will be priced in accord with the predictions of the model. Before turning to a formal analysis of this mechanism, we present some summary evidence regarding the distribution of relative ability by income. In Table 2, we present SAT math and verbal averages according to financial aid application status, which we use as a proxy for income. 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.

We also construct a relative ability measure for every student by taking the ratio of SAT math to SAT verbal. According to the 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, we use the tuition paid with the assumption that it is correlated with income. We 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

The empirical strategy in this section is based on the pricing equation derived from the model. According to Proposition 2, in the

Table 2
This table displays the means and standard deviations of SAT scores for groups according to financial aid application status.

	N	SAT math	SAT verbal	Non-academic factor
All	1161	717 (61)	645 (81)	2.54 (0.58)
F.A. applicants	796	710 (61)	654 (73)	2.62 (0.54)
F.A. recipients	743	711 (62)	655 (73)	2.62 (0.54)
Did not apply to F.A.	345	731 (57)	625 (93)	2.54 (0.65)

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

$$p_s^* \approx V'(k) + \eta^m(\theta^m - b_s^m) + \eta^v(\theta^v - b_s^v). \quad (12)$$

We assume tuition is measured with an additive error, and abilities are measured without errors. Then Eq. (12) can be written as:

$$p_s = \beta_0 + \beta_M \text{SATM}_s + \beta_V \text{SATV}_s + \varepsilon_s. \quad (13)$$

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. We use the grant amount g rather than tuition p , and focus on the financial aid applicants that decide to enroll.²⁶

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 we control for those too.²⁷ Finally, the financial aid is need based in principle. So we incorporate a need term to control for variation in income. The equation we estimate is:

$$g_s = \beta_0 + \beta_M \text{SATM}_s + \beta_V \text{SATV}_s + \beta_{NA} \text{NONACAD}_s + \beta_N \text{NEED}_s + \beta_F \text{FEMALE}_s + \beta_B \text{BLACK}_s + \beta_H \text{HISPANIC}_s + \varepsilon_s. \quad (14)$$

The model predicts nonnegative coefficients on all variables, and implies β_V to be strictly greater than β_M .

For some applicants, the aid package has a component called "reaction." 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" is that increment. Whether or not the reaction aid is offered and its amount 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 aid 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.²⁸ In both cases, the role of the reaction offer on enrollment cannot be ignored, so we present results with both the initial offer and offer with reaction used as the amount of aid, controlling for asking a reaction in the latter case.

²⁶ The property of interest, Prediction 1, is defined over enrolling students.

²⁷ Female and Minority Status differ from the characteristics we 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.

²⁸ A formal model of this multi-stage process determining financial aid is presented in Epple et al. (2006b).

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 Columns C and D provide the OLS and Tobit estimates respectively. The sample consists of enrolled financial aid applicants. Standard errors are given in parentheses.

	(A)	(B)	(C)	(D)
Intercept	-16,795 (3158)	-20,720 (3536)	-14,121 (3149)	-19,334 (3643)
SAT math	5.55 (3.58)	5.40 (3.98)	3.75 (3.56)	3.75 (4.09)
SAT verbal	12.48 (2.79)	14.57 (3.13)	11.61 (2.77)	14.80 (3.22)
Non-academic	2691 (364)	3066 (409)	2303 (362)	2696 (421)
Need	0.46 (0.01)	0.52 (0.02)	0.47 (0.01)	0.54 (0.02)
Female	1098 (415)	1056 (464)	1462 (413)	1519 (475)
Black/Native American	10,017 (717)	10,290 (790)	9561 (715)	9791 (810)
Hispanic	6952 (818)	7270 (901)	7075 (814)	7485 (922)
R ²	0.66	-	0.66	-
N	805	805	805	805
N-censored	-	125	-	125
Log-likelihood	-	-7195.25	-	-6959.80

Table 3 summarizes the main findings of the 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 the prediction about the pricing of abilities: The rare type of ability in a school is awarded more.²⁹

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).³⁰ This implies larger 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, relatively high math profile ($b^m > b^v$) is associated with higher-income students as indicated earlier. This implies that the direction of subsidies is from higher-income to lower-income types, as predicted.

Among other ability related variables, we find non-academic factor and high school rank to be important components of financial aid. The estimates for non-academic factor range from 2189 to 2899. We find no significant effect of high school GPA on financial aid. We interpret high school rank as a qualitative variable. We have tried different percentiles and found top 10% 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

²⁹ 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.

³⁰ The standard deviations are 83 and 67 after correcting for censoring at 800.

dependent variable includes the reaction gift, the estimates suggest that applicants who ask 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. Our estimates indicate that CMU is one of them. We find that African-Americans and Native-Americans receive a discount around ten thousand dollars. The discount is around seven thousand dollars for Hispanics. We do not find a significant discount for Asian or Other-Nonwhite categories. The estimates of the discount for women range from 854 to 1325 dollars.

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 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 our 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 multi-dimensional 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 researchers as well as 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. We leave these issues for future research.

Acknowledgements

I am indebted to Dennis Epple and Holger Sieg for their constant support. I would like to thank the co-editor Antonio Merlo, an anonymous referee, Richard Romano, Michael Peress, Gerhard Glomm, Insan Tunali, Maria Ferreyra, Bruce Weinberg, Michael Kaganovich, and Kenneth Smith, in addition to seminar participants at Carnegie Mellon University, Ohio State University, Indiana University, Koç University, EEA in Vienna, ESEM in Budapest, NASM at Duke, and APET in Nashville for their comments and suggestions. Any errors are mine. Thanks also to Bill Elliott, Vice President for Enrollment at

Carnegie Mellon, for the data used in this paper. Financial support from TÜBİTAK is gratefully acknowledged.

Appendix A

Proof of Proposition 1. Suppose $\theta_i^m = \theta_j^m$ and $\theta_i^y = \theta_j^y$ for some $i \neq j$ s.t. $i, j \neq 0$. Condition (7) implies $p_i(b^m, b^y, y) = p_j(b^m, b^y, y)$ for all (b^m, b^y, y) . Condition (8) and market clearance imply

$$EMC_i(b^m, b^y) = p_i(b^m, b^y, y) \geq p_j(b^m, b^y, y) = EMC_j(b^m, b^y)$$

for those students who attend school j , and

$$EMC_j(b^m, b^y) = p_j(b^m, b^y, y) \geq p_i(b^m, b^y, y) = EMC_i(b^m, b^y)$$

for those who attend school i . Linearity of EMC implies one of the following three cases: (a) $EMC_i(b^m, b^y) = EMC_j(b^m, b^y)$ for all (b^m, b^y) . (b) $EMC_i(b^m, b^y) > EMC_j(b^m, b^y)$ for all (b^m, b^y) (c) $EMC_i(b^m, b^y) \leq EMC_j(b^m, b^y)$ for some $c_1 b^m + c_2 b^y \geq \bar{c}$, for some constants c_1, c_2, \bar{c} . We 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:

$$p_i(b^m, b^y, y) = p_j(b^m, b^y, y) = EMC_i(b^m, b^y) = EMC_j(b^m, b^y) \tag{A1}$$

Now need to show that types (b_1^m, b_1^y, y_1) and (b_2^m, b_2^y, y_2) exist with $\alpha_i(b_1^m, b_1^y, y_1) \in (0, 1]$ and $\alpha_i(b_2^m, b_2^y, y_2) \in (0, 1]$ s.t. school i can increase profits by admitting the same number of (b_1^m, b_1^y, y_1) types as it expels (b_2^m, b_2^y, y_2) types. We have:

$$\begin{aligned} \pi_1^i &\equiv \frac{\partial \pi^i}{\partial [\alpha_i(b_1^m, b_1^y, y_1)] f(b_1^m, b_1^y, y_1)} \\ &= p_i^*(b_1^m, b_1^y, y, \theta_i^m, \theta_i^y) - V'(k_i) \\ &\quad + (b_1^m - \theta_i^m) \frac{1}{k_i} \iiint_S \frac{\partial p_i^*}{\partial \theta_i^m} \alpha_i f db^m db^y dy \\ &\quad + (b_1^y - \theta_i^y) \frac{1}{k_i} \iiint_S \frac{\partial p_i^*}{\partial \theta_i^y} \alpha_i f db^m db^y dy \end{aligned}$$

(π_2^i is analogous) and

$$\begin{aligned} \pi_{11}^i &\equiv \frac{\partial^2 \pi^i}{\partial^2 [\alpha_i(b_1^m, b_1^y, y_1)] f(b_1^m, b_1^y, y_1)} \\ &= -V''(k_i) \\ &\quad + 2 \frac{(b_1^m - \theta_i^m)}{k_i} \left[\frac{\partial p_i^*}{\partial \theta_i^m}(b_1^m, b_1^y, y_1, \theta_i^m, \theta_i^y) - \frac{1}{k_i} \iiint_S \frac{\partial p_i^*}{\partial \theta_i^m} \alpha_i f db^m db^y dy \right] \\ &\quad + 2 \frac{(b_1^y - \theta_i^y)}{k_i} \left[\frac{\partial p_i^*}{\partial \theta_i^y}(b_1^m, b_1^y, y_1, \theta_i^m, \theta_i^y) - \frac{1}{k_i} \iiint_S \frac{\partial p_i^*}{\partial \theta_i^y} \alpha_i f db^m db^y dy \right] \\ &\quad + \left(\frac{b_1^m - \theta_i^m}{k_i} \right)^2 \iiint_S \frac{\partial^2 p_i^*}{\partial^2 \theta_i^m} \alpha_i f db^m db^y dy \\ &\quad + \left(\frac{b_1^y - \theta_i^y}{k_i} \right)^2 \iiint_S \frac{\partial^2 p_i^*}{\partial^2 \theta_i^y} \alpha_i f db^m db^y dy \\ &\quad + 2 \frac{(b_1^m - \theta_i^m)}{k_i} \frac{(b_1^y - \theta_i^y)}{k_i} \iiint_S \frac{\partial^2 p_i^*}{\partial \theta_i^m \partial \theta_i^y} \alpha_i f db^m db^y dy \end{aligned}$$

(π_{22}^i is analogous) and

$$\begin{aligned} \pi_{12}^i &\equiv \frac{\partial \pi^i}{\partial [\alpha_i(b_1^m, b_1^v, y_1) f(b_1^m, b_1^v, y_1)] \partial [\alpha_i(b_2^m, b_2^v, y_2) f(b_2^m, b_2^v, y_2)]} \\ &= -V'(k_i) \\ &\quad + \frac{(b_2^m - \theta_i^m)}{k_i} \frac{\partial p_i^*(b_1^m, b_1^v, y, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} + \frac{(b_2^v - \theta_i^v)}{k_i} \frac{\partial p_i^*(b_1^m, b_1^v, y, \theta_i^m, \theta_i^v)}{\partial \theta_i^v} \\ &\quad - \left(\frac{b_1^m + b_2^m - 2\theta_i^m}{k_i}\right) \frac{1}{S} \iint \int \frac{\partial p_i^*}{\partial \theta_i^m} \alpha_i f db^m db^v dy \\ &\quad - \left(\frac{b_1^v + b_2^v - 2\theta_i^v}{k_i}\right) \frac{1}{S} \iint \int \frac{\partial p_i^*}{\partial \theta_i^v} \alpha_i f db^m db^v dy \\ &\quad + \frac{(b_1^m - \theta_i^m)}{k_i} \frac{\partial p_i^*(b_2^m, b_2^v, y_2, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} + \frac{(b_1^v - \theta_i^v)}{k_i} \frac{\partial p_i^*(b_2^m, b_2^v, y_2, \theta_i^m, \theta_i^v)}{\partial \theta_i^v} \\ &\quad + \frac{(b_1^m - \theta_i^m)}{k_i} \frac{(b_2^m - \theta_i^m)}{k_i} \iint \int \frac{\partial^2 p_i^*}{\partial^2 \theta_i^m} \alpha_i f db^m db^v dy \\ &\quad + \frac{(b_1^v - \theta_i^v)}{k_i} \frac{(b_2^v - \theta_i^v)}{k_i} \iint \int \frac{\partial^2 p_i^*}{\partial^2 \theta_i^v} \alpha_i f db^m db^v dy \\ &\quad + \left(\frac{(b_1^m - \theta_i^m)(b_2^v - \theta_i^v)}{k_i} + \frac{(b_1^v - \theta_i^v)(b_2^m - \theta_i^m)}{k_i}\right) \iint \int \frac{\partial^2 p_i^*}{\partial \theta_i^m \partial \theta_i^v} \alpha_i f db^m db^v dy \end{aligned}$$

Let Δ_i be the number of types (b_i^m, b_i^v, y_i) enrolled in school i . For Δ_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:

$$\pi_1^i \Delta_1 + \pi_2^i \Delta_2 + \frac{1}{2} [\pi_{11}^i (\Delta_1)^2 + \pi_{22}^i (\Delta_2)^2 + 2\pi_{12}^i \Delta_1 \Delta_2].$$

Since prices equal to effective marginal costs $\pi_1^i = \pi_2^i = 0$. We will consider admission changes such that $\Delta_1 = -\Delta_2$, so the above equation simplifies to

$$\Delta^2 \frac{1}{2} [\pi_{11}^i + \pi_{22}^i - 2\pi_{12}^i].$$

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

$$\begin{aligned} &= sign\left\{ \frac{2}{k_i} (b_1^m - b_2^m) \left[\frac{\partial p_i^*(b_1^m, b_1^v, y_1, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} - \frac{\partial p_i^*(b_2^m, b_2^v, y_2, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} \right] \right. \\ &\quad + \frac{2}{k_i} (b_1^v - b_2^v) \left[\frac{\partial p_i^*(b_1^m, b_1^v, y_1, \theta_i^m, \theta_i^v)}{\partial \theta_i^v} - \frac{\partial p_i^*(b_2^m, b_2^v, y_2, \theta_i^m, \theta_i^v)}{\partial \theta_i^v} \right] \\ &\quad + \frac{1}{k_i^2} (b_1^m - b_2^m)^2 \iint \int \frac{\partial^2 p_i^*}{\partial^2 \theta_i^m} \alpha_i f db^m db^v dy \\ &\quad + \frac{1}{k_i^2} (b_1^v - b_2^v)^2 \iint \int \frac{\partial^2 p_i^*}{\partial^2 \theta_i^v} \alpha_i f db^m db^v dy \\ &\quad \left. + \frac{2}{k_i^2} (b_1^m - b_2^m)(b_1^v - b_2^v) \iint \int \frac{\partial^2 p_i^*}{\partial \theta_i^m \partial \theta_i^v} \alpha_i f db^m db^v dy \right\} \end{aligned}$$

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

$$\begin{aligned} &= sign\left\{ \frac{2}{k_i} (b_1^m - b_2^m) \left[\frac{\partial p_i^*(b_1^m, b_1^v, y_1, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} - \frac{\partial p_i^*(b_2^m, b_2^v, y_2, \theta_i^m, \theta_i^v)}{\partial \theta_i^m} \right] \right. \\ &\quad \left. + \frac{1}{k_i^2} (b_1^m - b_2^m)^2 \iint \int \frac{\partial^2 p_i^*}{\partial^2 \theta_i^m} \alpha_i f db^m db^v dy \right\} \end{aligned}$$

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

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

Lemma 4 presented below shows that $p_i^*(b^m, b^v, y, \theta^m, \theta^v)$ is continuous in (b^m, b^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 $(b_1^m - b_2^m)$ and again using the same lemma.

Proof of Proposition 5. According to Proposition 1, θ_1^m cannot equal to θ_2^m since otherwise symmetry would imply two identical colleges. Let college 1 be the one with higher θ^m . Denote a student's achievement at College i by a_i . This symmetric equilibrium implies:

$$a_1 \geq a_2 \text{ when } b^m \geq b^v. \tag{B1}$$

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:

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

$$i, j \in \{0, 1, 2\}, i \neq j,$$

$$\text{and } \forall (b^m, b^v, y) EMC_0(b^m, b^v, y) = 0.$$

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:

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

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

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

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

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

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. $\eta_i^m = \eta_j^v \neq i$. If the shadow price of the rare ability is higher in a college, then $\theta_1^m > \theta_1^v$ implies:

$$EMC_1 \geq EMC_2 \text{ when } b^m \geq b^v. \tag{B3}$$

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 $\theta_1^m > \theta_1^v$ implies

$$EMC_1 \leq EMC_2 \text{ when } b^m \geq b^v \tag{B4}$$

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.

Lemma 1. If $\alpha_i(b, y) > 0$, then: $U(y - p_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) = U(y - EMC_k(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) \geq U(y - EMC_k(b^m, b^v), a(\theta_k^m, \theta_k^v, b^m, b^v))$ for some $j \neq i$ and for all $k \neq i$ s.t. $i, j, k = 1, 2, \dots, n$.

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 Eq. (3) for i implies $U(y - p_i(b^m, b^v, y), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_k(b^m, b^v, y), a(\theta_k^m, \theta_k^v, b^m, b^v))$ implying $p_k > EMC_k$ for this type. By Eq. (8), this implies $\alpha_k = 1$. Then, since $\alpha_i > 0$, market clearance is violated. This proves the inequality. Suppose $U(y - p_i(b^m, b^v, y), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - EMC_j(b^m, b^v), a(\theta_j^m, \theta_j^v, b^m, b^v))$ 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.

$$U^*(b^m, b^v, y) = \max_{i \in \{1, 2, \dots, n\} | \text{college } i \text{ is not attended by } (b^m, b^v, y)} U(y - EMC_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) \text{ for all } (b^m, b^v, y).$$

Proof. Follows immediately from Lemma 1.

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

Proof. For any positive measure in type space S of students, $U^*(b^m, b^v, y)$ is the upper envelope of a set of continuous functions in (b^m, b^v, y) by the above lemma. Hence $U^*(b^m, b^v, y)$ 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 $U(y - p_i(b^m, b^v, y), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_j(b^m, b^v, y), a(\theta_j^m, \theta_j^v, b^m, b^v))$. Since the student attends college i , $p_i(b^m, b^v, y) > EMC_i(b^m, b^v)$, and $p_j(b^m, b^v, y) < EMC_j(b^m, b^v)$ or $\alpha_j(b^m, b^v, y) = 1$ and market does not clear. Then $U(y - EMC_i(b^m, b^v), a(\theta_i^m, \theta_i^v, b^m, b^v)) > U(y - p_j(b^m, b^v, y), a(\theta_j^m, \theta_j^v, b^m, b^v))$. 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, y) .

Proof. This is implied by Lemma 3 and Eq. (7).

References

- Acemoglu, D., 2002. Technical change, inequality, and the labor market. *Journal of Economic Literature* 40–1, 7–72.
- Arcidiacono, P., Nicholson, S., 2005. Peer effects in medical school. *Journal of Public Economics* 89 (2–3), 327–350.
- Battistin, E., Blundell, R., Lewbel, A., 2009. Why is consumption more log normal than income? Gibrat's Law revisited. *Journal of Political Economy* 117 (6), 1140–1154.
- Benabou, R., 1996. Heterogeneity, stratification, and growth: macroeconomic effects of community structure. *American Economic Review* 86, 584–609.
- Benabou, R., 2002. Tax and education policy in a heterogeneous-agent economy: what levels of redistribution maximize growth and efficiency? *Econometrica* 70, 481–517.

- Blanden, J., Gregg, P., Macmillan, L., 2007. Accounting for intergenerational income persistence: noncognitive skills, ability and education. *The Economic Journal* 117, 43–60.
- Bowles, S., Gintis, H., Osborne, M., 2001. The determinants of earnings: a behavioral approach. *Journal of Economic Literature* 39, 1137–1176.
- Cautt, E., 2002. Educational policy when there are peer group effects – size matters. *International Economic Review* 43, 195–222.
- Cunha, F., Heckman, J., 2007. The technology of skill formation. *American Economic Review* 97 (2), 31–47.
- Cunha, F., Heckman, J., Lochner, L., Masterov, D., 2006. Interpreting the evidence on life cycle skill formation. In: Hanushek, E., Welch, F. (Eds.), *Handbook of the Economics of Education*. North Holland.
- Dale, S., Krueger, A., 2002. Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables. *Quarterly Journal of Economics* 117 (4), 1491–1527.
- deBartolome, C., 1990. Equilibrium and inefficiency in a community model with peer group effects. *Journal of Political Economy* 98 (1), 110–133.
- Dragulescu, A., Yakovenko, V.M., 2001. Evidence for the exponential distribution of income in the U.S.A. *The European Physical Journal B* 20, 585–589.
- Epple, D., Romano, R., 1998. Competition between private and public schools, vouchers and peer group effects. *American Economic Review* 88, 33–63.
- Epple, D., Romano, R., Sieg, H., 2003. Peer effects, financial aid, and selection of students into colleges and universities: an empirical analysis. *Journal of Applied Econometrics* 18 (5), 501–526.
- Epple, D., Romano, R., Sieg, H., 2006a. Admission, tuition, and financial aid policies in the market for higher education. *Econometrica* 74, 885–928.
- Epple, D., Romano, R., Sarpa, S., Sieg, H., 2006b. Profiling in bargaining over college tuition. *The Economic Journal* 116–515, 459–479.
- Fernandez, R., Rogerson, R., 1998. Public education and income distribution: a dynamic quantitative evaluation and education finance reform. *American Economic Review* 88, 813–833.
- Fernandez, R., Rogerson, R., 2003. Equity and resources: an analysis of education finance systems. *Journal of Political Economy* 111, 858–897.
- Grogger, J., Eide, E., 1995. Changes in college skills and the rise in the college wage premium. *The Journal of Human Resources* 30 (2), 280–310.
- Hamermesh, D.S., Donald, S.G., 2006. The effect of college curriculum on earnings: an affinity identifier for non-ignorable non-response bias. Working Paper.
- Henderson, V., Mieszkowski, P., Sauvageau, Y., 1978. Peer group effects and educational production functions. *Journal of Public Economics* 10, 97–106.
- Hoxby, C., Weingarth, G., 2005. Taking race out of the equation: school reassignment and the structure of peers effects. Working Paper.
- James, E., Alsalam, N., Conaty, J., To, D., 1989. College quality and future earnings: where should you send your child to college? AEA Papers and Proceedings 79, 247–252.
- Koenig, K., Frey, M., Detterman, D., 2008. ACT and general cognitive ability. *Intelligence* 36, 153–160.
- Manski, C., 1993. Identification of endogeneous social effects: the reflection problem. *Review of Economic Studies* 60, 531–542.
- Nechyba, T., 2000. Mobility, targeting and private-school vouchers. *American Economic Review* 90, 130–146.
- Paglin, M., Rufolo, A.M., 1990. Heterogeneous human capital, occupational choice, and male–female earnings differences. *Journal of Labor Economics* 8 (1), 123–144.
- Roy, A.D., 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3 (2), 135–146.
- Sacerdote, B., 2001. Peer effects with random assignment: results for Dartmouth roommates. *Quarterly Journal of Economics* 116 (2), 681–704.
- Scotchmer, S., 1994. Public goods and the invisible hand. In: Quigley, J., Smolensky, E. (Eds.), *Modern Public Finance*. Harvard University Press.
- Stinebrickner, T., Stinebrickner, R., 2006. What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics* 90, 1435–2454.
- Winston, G., Zimmerman, D., 2004. Peer effects in higher education. In: Hoxby, C. (Ed.), *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*. The University of Chicago Press.
- Zimmerman, D., 2003. Peer effects in academic outcomes: evidence from a natural experiment. Williamstown, MA, The Williams Project on the Economics of Higher Education. *Review of Economics and Statistics* 85 (1), 9–23.