Market Power and Price Discrimination in the U.S. Market for Higher Education*

Dennis Epple

Carnegie Mellon University

Richard Romano

University of Florida

Sinan Sarpça

Koç University

Holger Sieg

University of Pennsylvania

Melanie Zaber

RAND Corporation

We estimate an equilibrium model of private and state college competition that generates realistic pricing patterns for private colleges using a large national data set from the NPSAS. Our analysis distinguishes between tuition variation that reflects efficient pricing to students who generate beneficial peer externalities and variation that reflects arguably inefficient exercise of market power. Our findings indicate substantial exercise of market power and, importantly, sizable variation in this power along the college quality hierarchy and among students with different characteristics. Finally, we conduct policy analysis to examine the consequences of increased availability of quality public colleges in a state.

*Epple: epple@cmu.edu; Romano: romanor@ufl.edu, Sarpca: ssarpca@ku.edu.tr, Sieg: holgers@econ.upenn.edu: Zaber: mzaber@rand.org. The authors thank two anonymous referees, Philipp Kircher, Charles Manski, Derek Neal, Cecilia Rouse, Marc Rysman, Petra Todd, Ken Wolpin for comments and discussions. We also thank the NSF for financial support (NSF SES-1355892). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Center for Education Statistics.

1 Introduction

The net tuition paid by students sitting in the same college classroom is often quite different, varying by student characteristics such as ability, income, and minority status.[1] Some variation in observed tuitions reflects efficient pricing to students that provide positive externalities to their classmates. Other variation reflects arguably inefficient exercise of market power. The main objective of this article is to quantify the magnitude of these different effects. To accomplish this goal we estimate an equilibrium model of private and public college competition in which private college pricing reflects a combination of exercise of market power and discounts to students who provide valuable peer externalities. The intuition for the latter is straightforward. Competition for students who provide desirable peer externalities induces colleges to offer discounts (financial aid) to attract higher ability students and students who increase diversity. The exercise of market power relies on a more subtle foundation. Idiosyncratic variation in student preferences permits colleges to charge more to high-income households. Colleges do not observe an individual student's idiosyncratic preferences. They know that some members of the applicant pool will have strong idiosyncratic preference for their school, and they know that high-income parents will pay a premium to accommodate preferences of their children. Hence, colleges charge a premium to high-income applicants, correctly anticipating that high-income applicants with high idiosyncratic preference for the college will select into the college. Colleges then use these additional revenues to enhance quality by cross-subsidizing low-income high-ability students and by increasing instructional expenditures.

Our empirical analysis builds on the model developed in Epple, Romano, Sarpca, and Sieg (2017). Private colleges choose admission, tuition, and expenditure policies to maximize a quality index, whereas state colleges choose admission policies and expenditure to maximize aggregate achievement of their in-state students facing state

[1] By net tuition, we mean tuition net of the student's institutional financial aid. We use posted tuition to refer to the amount paid by students that receive no institutional aid.

regulated tuitions. [2] Demand for colleges is modeled using a discrete choice random utility framework. The demand model accounts for the fact that not all students are admitted to selective colleges. Both private and state colleges optimally will use minimum ability admission thresholds. Given these admission thresholds, we can determine the set of colleges that are feasible for each student type.

The optimal financial aid or pricing policies of private schools have a number of interesting properties. First, because private schools set a posted or maximum tuition, a minimum ability threshold that characterizes admission policies of private colleges arises. A certain fraction of students do not obtain financial aid and pay this maximum tuition. These are students that are of relatively low ability in the school and thus do not qualify for merit aid. Moreover, these students must have income sufficiently high so that they are willing to pay the price maximum. Second, for other students, net tuition can be expressed as a convex combination of "effective marginal cost" and a mark-up on income.

Important differences distinguish this model from a standard oligopolistic pricing model. First, effective marginal cost depends on the ability and minority status of a student. Pricing by ability or merit-based aid arises because high ability students increase college quality through peer and reputational effects. Discounts for minority students arise because they enhance diversity.[3] Second, the mark-up term for a student does not depend on the overall market share of the college, but on the market share conditional on observed student characteristics. We show that these terms can differ by large margins among students, especially for highly selective colleges. Hence "local" or conditional market shares drive mark-ups in the model, and not overall market shares. Third, the mark-up is monotonically increasing in student or household income and arises due to price discrimination. As a consequence, this model

[2]The private college objective can be interpreted as a desire to maximize their reputation. The best objective to attribute to colleges is an open question, and we discuss alternatives below.

[3]Our article is, therefore, also related to the recent literature that estimates empirical models of affirmative action. Some recent work includes Hickman (2013), Kapor (2016), and Cestau, Epple, and Sieg (2017).

is sufficiently rich to generate the qualitative features of tuition policies observed in the U.S. market for higher education.

The main contribution of this article is that we derive and implement a new semi-parametric estimator for the parameters of the model. We can identify and estimate almost all parameters of the model using a method of moments estimator that is based on the difference between the observed and predicted price functions at private colleges. To implement this estimator we need a non-parametric plug-in estimator of the conditional market share for each student at the school that is attended in equilibrium. Our estimation approach does not require us to solve for the equilibrium of the model. [4] This has the virtue of simplicity. It is also computationally feasible as we do not need to use a nested fixed point algorithm.[5] Moreover, the estimator is consistent for a finite number of elements in the choice set as the number of students in a single cross section goes to infinity.

We implement this estimator using data from the National Postsecondary Student Aid Study (NPSAS). Our sample size consists of approximately 9,500 students that attended a two-year public community college, a four-year public college, or a four-year private college in the U.S. in 2012. Approximately 2,270 students attend a four-year private college. This subsample is the core sample used in estimation. The remaining students that attend public schools are primarily used to estimate the conditional market shares.

Our sample size is large. It is, however, not large enough to estimate a model at the individual college level.[6] We, therefore, use clustering algorithms to aggregate four-year private colleges into ten types, public four-year colleges into four types, and public two-year colleges into one type. Our empirical demand model thus has 15

[4]The idea of conditioning on observed choice probabilities is similar to Heckman (1979), Berry (1994), and Berry, Levinsohn, and Pakes (1995).

[5]Our quasi-maximum likelihood estimator is most similar to the one proposed by Bajari, Hong, and Nekipelov (2010) to estimate games with incomplete information.

[6]The NPSAS is the most comprehensive data set available for the U.S., but only samples a subset of all colleges in the U.S. As a consequence some sort of aggregation is unavoidable if one estimates any demand model for higher education using this data set.

different college types.

We find that the majority of private colleges engage in pricing by income, ability, and minority status. A $10,000 increase in family income increases tuition at private schools by on average $210 to $510. A one standard deviation increase in ability decreases tuition by approximately $920 to $1,960 depending on the selectivity of the college. Large and substantial discounts for minority students arise that range between approximately $110 (at historically black colleges) and $5,750. Average mark-ups in private colleges range between $750 and $13,200. Much more heterogeneity and some much larger mark-ups occur within colleges than the averages.

Finally, we conduct some policy experiments. We show that there are significant gains from having access to a diversified set of public schools. There are direct effects for students attending these schools. In addition, there are indirect effects for students that chose to attend private schools in equilibrium as they benefit from the fact that they have better outside options. Moving students from a state with a low quality public education system to a state which has a fully diversified public school system increases average welfare by approximately $465 per student (among students affected by the move).

Our article is related to at least three different areas of research that have focused on markets for higher education. Most importantly, other articles have documented that pricing by income, ability and minority status is prevalent in the financial aid data (e.g., McPherson and Schapiro (2006)).[7] Most of this empirical work is nonstructural.[8] Relative to this literature, our contribution is to estimate a structural model that provides a detailed explanation as to how and why this pricing arises and varies among and within colleges. Two other structural articles tackle these issues. Epple, Romano, and Sieg (2006) estimate a model of competitive college pricing with infor-

[7] van der Klaauw (2002) uses a regression discontinuity strategy to confirm that colleges use price discounts to attract students.

[8] For a discussion of that literature see, among others, McPherson and Schapiro (1998), Ehrenberg (2002) , Dynarski (2002), Epple, Romano, and Sieg (2003), Bettinger, Long, Oreopolous, and Sanbonmatsu (2005), Fillmore (2015), and Sieg and Wang (2018).

mational assumptions that imply first-degree price discrimination and also assumes colleges provide need-based aid to foster income diversity. Important differences in the present analysis are: (i) unobserved idiosyncratic student preferences; (ii) state college competition; (iii) the inclusion of race and state of origin as observed heterogeneity; (iv) a new estimation strategy; and (v) use of a more extensive data set. Fillmore (2015) also studies price discrimination in an alternative model of college competition and explores restricting colleges' ability to use some or all of the information incorporated in the Free Application for Federal Student Aid (FAFSA). Our model differs substantially, most importantly having variation in college qualities that arises endogenously, with, again, a rich set of predictions about variation in college pricing.[9]

Second, our work is related to research that has modeled admission and attendance decisions in the market for higher education. The informational environment in our model implies students face no uncertainty in admissions, so we can abstract from an application-admission game with incomplete information. Avery and Levin (2010), Chade, Lewis, and Smith (2014) and Fu (2014) provide a detailed analysis of these issues. Our model also abstracts from choices made by students once they enter college. The most important decision is the choice of a major. Arcidiacono (2005) and Bordon and Fu (2015) develop and estimate dynamic models of choice of academic major under uncertainty. Last, in Epple et al. (2017) , we developed the basic theoretical model to examine effects of policy changes on attendance patterns and student costs. We calibrated the parameters of the model and imposed an equilibrium selection criteria. These simplifications allowed us to numerically solve for equilibrium in the model and perform policy analysis. In this article we extend the model to allow for differences in minority status. Most importantly, we develop and implement a semi-parametric estimator for the parameters of the model. Moreover,

[9]As our contribution is empirical, we largely avoid discussion of theoretical models of college competition. An exception we note is Rothschild and White (1995), who are the first to provide a model that shows pricing of within college student externalities can arise in a competitive setting.

we conduct an empirical analysis of price discrimination and market power here.

Finally, our article is related to recent research on the importance of peer effects in education. Epple and Romano (2010) and Sacerdote (2011) provide a literature surveys. We do not provide any direct evidence on the importance of peer effects, but provide strong indirect evidence based on our analysis of pricing by ability and minority status.

The rest of the article is organized as follows. Section 2 introduces our data set and provides descriptive statistics. Section 3 introduces the model that characterizes student sorting among and price and admission policies of colleges. Section 4 introduces a parametrization and discusses our estimator. Section 5 reports our parameter estimates. Section 6 explores the implications of this analysis for mark-ups, market power and price discrimination. Section 7 discusses some of the policy implications of our results. Section 8 concludes the analysis.

2 Data

Our data source is the 2011-12 National Postsecondary Student Aid Study (NPSAS) from the National Center for Education Statistics (NCES).[10] Our model focuses on initial attendance/matriculation outcomes. We construct our sample using first-year students, who are oversampled in this wave of the NPSAS and constitute more than half of all observations.[11]

We construct our sample to reflect the pool of students who would plausibly consider four-year private schools. Students who would only ever consider a two-year institution are outside the scope of this article; thus, we drop students who have what would be "atypical attendance" patterns at a four-year institutions. This in-

[10] The NPSAS data are accompanied by inverse probability weights that account for the composite probability of sampling, both at the college and individual level. We use these weights throughout the empirical analysis.

[11] College completion and continuation decisions are likely to differ from the initial matriculation decision. Also, family resources and aid packages in later years of attendance need not be identical to those in the student's first year. For these reasons we use first-year students in our analysis.

cludes multiple attenders–students who switch institutions in their first academic year.[12] This also includes a number of students who attend part-time or part-year. This eliminates the majority of two-year students. These students are unlikely to have considered four-year colleges at the time of application: only five percent of exclusively part-time two-year college students from the 2010 entry cohort ever transferred to a four-year institution, compared to 31.5 percent overall (National Student Clearinghouse 2017).[13]

We also drop veterans and athletes because their financial aid opportunities are different from those faced by the average student, and their priorities in selecting an institution may also differ. We drop foreign students (or students with no state residence) for two reasons: (i) Their choice sets possibly include the universities in their home country, as well as universities in other non-home countries (based on their decision to study abroad); and (ii) their eligibility for financial aid and their pricing by colleges may differ.

Ability is a key variable in our analysis, and we drop observations with missing components of the ability measures (ACT or SAT score and high school GPA).[14] We drop all students attending schools at which we cannot match institutional expenditures. Finally there are a few sample schools that offer both 4-year and 2-year degrees, and we drop their 2-year enrollees (the minority) and treat them as 4-year institutions. The resulting sample consists of approximately 9,490 students. Table 1 presents the numbers of these groups of students along with their distribution over different types of colleges.

Table 2 presents selected statistics from our sample. Our measure of ability is

[12]These constitute about 4 percent of the sample, dropped because we cannot know if the switch was planned from the point of matriculation, and so the decision space would become much more complex.

[13]We recognize that transferring to a four-year institution is not the same decision process as having considered a four-year institution at the time of application; however, the latter counterfactual is unobservable in data.

[14]As approximately 40% of first-year public 2-year students do not take SATs or ACTs, it is possible that the remaining sample is of higher ability than the general student body. Thus, our measure of average peer quality may be biased upward for the lowest tier college.

predicted college GPA–we model college GPA as a function of high school GPA, ACT or SAT score, gender, major, and college fixed effects in a sample of non-minority four-year college students. We then predict GPA at a generic college, using only the recovered parameters for high school GPA, ACT/SAT score, and gender.[15] This ability measure is then transformed to have unit standard deviation and positive mean. The choice of mean ensures that the average ability at each college is weakly greater than zero.[16]

Our measure of income is adjusted gross income in 2010. Where possible, NPSAS computes this value based on the federal financial aid application, and uses total income (of family or student as implied by dependency status) reported in the student interview where no application or tax return are available. The 2010 value is used as federal financial aid eligibility for 2011-2012 school year would be based on 2010 income. Race, ethnicity, and gender are drawn from the student interview where possible, and from student records when no interview is possible.

In-state status is determined by comparing the student's reported state of residence with the imputed availability of public college types.[17] We calculate total institutional aid by taking the sum of institutional grants, one-half of work study, and one-quarter of loans. Thus, net tuition is the posted tuition less the sum of institutional aid (federal aid is considered separately).

Federal aid is limited to Pell grants, which are calculated by the formula

$$A = \min\Big\{\max\{0, COA - EFC(y)\}, 5500\Big\}, \tag{1}$$

where COA is the federally determined cost–of-attendance and $EFC(y)$ the federally determined expected family contribution, which increases with household income.

[15] We do not account for minority status in this regression although it could be easily done. A priori one can make arguments in favor and against either approach.

[16] Appendix A provides additional details. It also explains the construction of the ability thresholds used for each cluster.

[17] The imputation procedure is discussed in Appendix B.

Pell grants are awarded up to $COA - EFC$ if positive, but with a maximum of \$5500. However, in practice, we use the amended formula:

$$A = \min\Big\{\max\{0, \bar{p_j} + L - EFC(y)\}, 5500\Big\}, \tag{2}$$

as cost of attendance – posted tuition ($\bar{p_j}$) plus estimated non-tuition costs (L) – varies by student-college combination and is only occasionally observed at the attended college, and never observed for potential alternatives. EFC is directly reported in the NPSAS, and thus can be used both for the attended college as well as the potential alternatives.

Then we calculate the Pell aid at each college using the above formula, also adjusting to account for the Pell minimum award (in 2012, 555 dollars). Any student offered at least half of the minimum, but less than the minimum, is given the minimum, and any student eligible for less than half of the minimum was awarded no aid. Additionally, we have many "never-takers" in our sample, and so if we observe a student to be a non-taker at the attended college when eligible for some aid, we assume the student is a never-taker at all colleges. The resulting data are then merged by cluster to data on instructional expenditures from the Delta Cost Project Database (National Center for Education Statistics) for academic year 2012.

Our sample includes observations from approximately 900 colleges. The number of students observed per college averages about 11. Having more observations per college is desirable for precision when testing within-college predictions of the model. At the same time, our model implies that colleges with similar characteristics would make similar admission and pricing decisions. Working with smaller choice sets (fewer colleges) also has computational advantages. For these reasons, we group together colleges that are similar in their key characteristics, treating them as one college. We group public and private colleges separately based on the joint variance of posted tuition, average ACT score, and instructional expenditures per student, using k-means clustering. We choose the number of clusters based on the elbow method (Goutte, Toftand, Rostrup, Nielsen, and Hansen, 1999), increasing the number of clusters until the marginal cluster does not significantly decrease the within-group variance, which suggests approximately four clusters of public four-year colleges, and approximately twelve clusters of private four-year colleges.

The "rule of thumb" relates the suggested k to the number of schools to cluster, $k = \sqrt{\frac{n}{2}}$, implies approximately 13 private clusters (Mardia, Kent, and Bibby, 1979). We initially create twelve private clusters, but then combine two sets of resulting cluster pairs to ensure an adequate sample size of students at each cluster. Table 3 presents the key characteristics of private and public clusters, ordered within the two college groups by mean ACT. The term "college" will refer to a cluster in the rest of the article.[18]

Columns (2) and (3) of Table 3 give the admission thresholds that we observe among non-legacy students.[19] We find that these thresholds are the same for private schools as we impose no residency-based restriction on those schools' admission, but they vary for public schools.

Grouping or clustering colleges implies that we assume that colleges within a group do not compete against each other and use the same pricing and admission rules. This assumption is important, but also unavoidable as we need to observe a large enough sample for each college cluster to estimate the model as discussed in detail below.

3 A Model of Price Discrimination

In this section, we briefly review the model derived in Epple, Romano, Sarpca, and Sieg (2017) and derive the key implications for optimal admission and pricing policies. Colleges vary by quality endogenously. The main difference between the model that

[18]Treating clusters as one college assumes: (i) students perceive colleges within a cluster as equivalent; and (ii) either students apply to no more than one college within a cluster or that colleges within clusters effectively collude. As we have 14 four-year clusters and students that qualify for a cluster can attend at their price, we are not significantly limiting the choice set of students.

[19]We eliminate a small number of observations with very low ability from the top two tiers as these students are likely to be legacy students. Our model does not capture these considerations.

we estimate in this article and the model that was used in Epple, Romano, Sarpca, and Sieg (2017) is that we add minority status as a student characteristic that affects quality, and thus study the pricing implications of preferences for diversity.

We consider a model with S regions or states and normalize the student population in the economy to 1. Let π_s denote the student population proportions or size of each state and note that $\sum_{s=1}^{S} \pi_s = 1$. Students in each state differ continuously by after-tax income y and ability b. Students also differ by minority status which is a discrete indicator variable $m \in \{0, 1\}$. Let $f_s(b, y|m)$ denote the density of (b, y) in state s conditional on m. The fraction of of type m households in state s is denoted by π_{sm}; note that $\sum_m \pi_{sm} = \pi_s$.

For expositional simplicity, we assume each state operates one public university. In our application discussed below, we extend the model and allow for quality differentiation among multiple public colleges within a state.[20] In addition to the S public universities, there are P private universities that operate nationwide and also compete for students. There is an outside option which we model as attending a two-year public college. The total number of alternatives is then $J = S + P + 1$.[21]

We next describe the demand side of the model and then turn to the supply side. A student with ability b that attends a university of quality q_j has an achievement denoted by $a(q_j, b)$. Let $p_{sj}(m, b, y)$ denote the tuition that a student from state s with ability b, income y , and minority status m pays for attending college j. Let $A_{sj}(y)$ denote federal aid and L the non-tuition cost of attending a college. Federal aid depends on income and the cost of attending a college, which varies with a student's state of residence if attending a state college. Let ε_j denote an idiosyncratic preference shock for college j, which is private information of the student.

The utility of student (s, m, b, y) attending college j is additively separable in the

[20]Our empirical model allows for up to four public college types in each state.

[21]We abuse notation for convenience by using S to denote both the number of state colleges and the set of them $\{1, 2, \ldots, S\}$, and likewise for P and J (which usage will be obvious by context).

idiosyncratic component and given by:

$$U_j(s,m,b,y,\varepsilon_j) \;=\; U(y - p_{sj}(m,b,y) - L + A_{sj}(y), a(q_j,b)) \;+\; \varepsilon_j. \tag{3}$$

$U(\cdot)$ is an increasing, twice differentiable, and quasi-concave function of the numeraire and educational achievement, $a(\cdot)$. Educational achievement is an increasing, twice differentiable, and strictly quasi-concave function of college quality and own ability.

Utility depends on location and minority status because tuition depends on location and minority status. The dependence on location can arise for two reasons. First, state colleges are likely to give preferential treatment to in-state students. Second, private colleges may use different mark-ups to students coming from different states because these students may face different state college options. The dependence on minority status follows from the fact that diversity affects quality and college pricing as discussed below.

Let $S_a(s,m,b)$ denote the subset of state colleges to which student (s,m,b,y) is admitted, $P_a(s,m,b)$ the same for private colleges, and $J_a(s,m,b) \subset S_a(s,m,b) \cup P_a(s,m,b) \cup O$ the options that provide positive utility available to the student.[22] Taking as given tuitions, qualities, and non-institutional aid, student (s,m,b,y) chooses among $j \in J_a(s,m,b)$ to maximize utility (3). Let the optimal decision rule be denoted by $\delta(s,m,b,y,\varepsilon)$. The vector ε satisfies standard regularity assumptions in McFadden (1973).

Integrating out the idiosyncratic taste components yields conditional choice probabilities for each type:

$$r_{sj}(m,b,y;P(m,s,b,y),Q) \;=\; \int 1\{\delta_j(s,m,b,y,\varepsilon)=1\}\; g(\varepsilon)\; d\varepsilon, \tag{4}$$

where $1\{\cdot\}$ is an indicator function, $\delta_j(\cdot)=1$ means college j is chosen, $P(s,m,b,y)$ denotes the vector of tuitions that apply to student type (s,m,b,y), and Q denotes

[22] Admission will be independent of income, though net tuition will vary with income.

the vector of college qualities.

Private colleges attract students from all states of the country. Their objective is to maximize quality constrained by demand, competition, and costs. College j has a cost function

$$C_j(k_j, I_j) = F_j + V_j(k_j) + k_j \; I_j, \tag{5}$$

where k_j denotes the size of college j's student body and I_j expenditures per student on educational resources in college j. The costs $F_j + V_j(k_j)$ are independent of educational quality. Each college also obtains an exogenous amount of non-tuition income denoted by E_j, e.g., from endowment earnings. Finally, private colleges also have exogenous posted or maximum tuitions, denoted by $\bar{p}_j$.[23]

Letting θ_j denote mean ability in college j's student body and Γ_j the fraction of minority students, college quality is given by

$$q_j = q_j(\theta_j, I_j, \Gamma_j) \tag{6}$$

which is a twice differentiable, increasing, and strictly quasi-concave function of $(\theta_j, I_j, \Gamma_j)$. Quality increases with average student ability due to a combination of peer learning effects, non-learning externalities from developing relationships with high ability peers, and reputation effects.[24] A key argument for the assumption that quality increases with diversity is that having diverse student peers may enhance post-college success in a diverse workplace. This assumes that minorities are under-represented in each university.[25] The reasons why quality is impacted by student

[23]Setting a maximum tuition, rather than having tuition strictly monotonic in income, is likely explained by a combination of social/political marketing and uninformed students who are unaware of financial aid. For tractability, we treat a private college's posted tuition as given in the theory.

[24]In as much as there are non-learning and reputation effects of having higher ability peers embodied in θ, what we have labeled "achievement" must be more broadly interpreted as any utility enhancing college effect. See, for example, MacLeod and Urquiola (2015).

[25]Research has also shown that high-achieving low-income students have poor information about college characteristics, including availability of aid they would receive (Hoxby and Turner, 2015). This can deter applying to selective colleges especially. This is not, however, an element of our model.

abilities and diversity are not relevant to the positive predictions we derive and test.

Private colleges maximize quality behaving as monopolistic competitors. Private college j takes as given other colleges' tuitions, admissions, and qualities when maximizing quality. The fact that private colleges take qualities of other colleges as given when choosing their admission, tuition, and expenditure policies implies the equilibrium is not Nash in these policies; rather, it implies a version of monopolistic competition. Private colleges ignore that varying their own choices would affect the qualities of other colleges. For example, a college policy change that attracts high ability students away from other colleges would lower the laters' qualities.[26] Private college j chooses price and admission functions, respectively, $p_{sj}(m,b,y)$ and $a_{sj}(m,b,y) \in [0,1]$, per student resource expenditure, I_j , and $(\theta_j, \Gamma_j, k_j)$ through the constraints to solve:

$$\max \ q(\theta_j, I_j, \Gamma_j) \tag{7}$$

subject to a revenue constraint

$$R_j = \int\int \sum_{s=1}^{S} \sum_{m} \pi_{sm} \ p_{sj}(m,b,y) r_{sj}(m,b,y; P(m,s,b,y), Q) a_{sj}(m,b,y) f_s(b,y|m) \ db \ dy + E_j \tag{8}$$

a budget constraint

$$R_j = F_j \ + \ V_j(k_j) \ + \ k_j I_j \tag{9}$$

identity constraints,

$$\theta_j \ = \ \frac{1}{k_j} \iint b \left(\sum_{s=1}^{S} \sum_{m} \pi_{sm} r_{sj}(m,b,y; P(m,s,b,y), Q) a_{sj}(m,b,y) f_s(b,y|m) \right) db \ dy \tag{10}$$

$$k_j \ = \ \iint \left(\sum_{s=1}^{S} \sum_{m} \pi_{sm} r_{sj}(m,b,y; P(m,s,b,y), Q) a_{sj}(m,b,y) f_s(b,y|m) \right) db \ dy, \tag{11}$$

[26] If individual colleges are not too big, then this assumption is reasonable. It also avoids potential existence problems.

$$\Gamma_j = \iint \left(\sum_{s=1}^{S} \pi_{s1} r_{sj}(1,b,y;P(1,s,b,y),Q) a_{sj}(1,b,y) f_s(b,y|1) \right) db\, dy \Big/ k_j, \quad (12)$$

and the maximum price constraint

$$p_{sj}(m,b,y) \leq \bar{p}_j. \quad (13)$$

We solve the private college's problem, with more detail provided in an appendix (available from the authors upon request).[27] Focusing on tuition, assuming that (11) is not binding, for any student (s,m,b,y) with $r_{sj} > 0$, tuition satisfies:

$$p_{sj}(m,b,y) + \frac{r_{sj}(m,b,y;\cdot)}{\partial r_{sj}(m,b,y;\cdot)/\partial p_{sj}(m,b,y)} = EMC_j(m,b) \quad (14)$$

where

$$EMC_j(m,b) \equiv V_j' + I_j + \frac{q_\theta}{q_I}(\theta_j - b) + \frac{q_\Gamma}{q_I}(\Gamma_j - m) \quad (15)$$

The left-hand side of (14) is the usual expression for marginal revenue. The right-hand side of expression (14) is the "effective marginal cost" of student (s,m,b,y)'s attendance, which sums the marginal resource cost given by the first two terms and the marginal peer costs given by the last two terms. The ability-based marginal peer cost (third term) multiplies the negative of the student's effect on the peer measure (equal to $(\theta - b)/k$) by the resource cost of maintaining quality (equal to $\frac{\partial q/\partial \theta}{\partial q/\partial I} k$). The diversity marginal peer cost has analogous decomposition. Note that EMC varies by student within college j only with the student's ability and minority status. The ability-based marginal peer cost is negative for students of ability exceeding the college's mean, and the diversity-based marginal peer cost is negative for minorities. Note, as well, that it is through the peer cost terms that optimal pricing to all types is intertwined. For example, lowering tuition to attract more high ability students

[27] The first-order conditions on $(\theta_j, \Gamma_j, I_j)$ are not presented here but are used to write the pricing equation on which we focus presented below. Of course, these variables are chosen to maximize quality. All of a student types are admitted ($a_{sj}(m,b,y) = 1$) unless the maximum price constraint binds, as further discussed below.

increases θ_j, implying higher prices are optimal to all types.

Because tuition increases to a type reduces attendance if preferred tuition is below the posted price, the optimal admission policy is simple. All students (s, m, b, y) are admitted to the college if and only if

$$\min\{\bar{p}_j, p_{sj}(m, b, y)\} \geq EMC_j(m, b) \tag{16}$$

for p_{sj} satisfying (12). All students of the type are admitted if (11) does not bind. Otherwise, equation (16) yields minimum ability thresholds that vary with minority status for each private college implicitly defined by:

$$\bar{p}_j = EMC_j(m, b_{jm}^{min}) \tag{17}$$

Effective marginal cost decreases with ability and is lower for a minority student of given ability. Hence, the admission threshold for minorities is lower.

Our assumption that private colleges maximize a quality index is arguably well motivated given colleges' preoccupation with rankings. The specification, especially when we add more structure in Section 4.1, implies a clear and tractable pricing structure that accords well with stylized facts about college pricing, namely merit and need-based aid and discounting to under-represented minorities. It facilitates testing, and we find significance of the key underlying parameters. However, the proper college objective and the role of diversity in it is an open question. Colleges might, for example, discount to minorities because their utilities directly enter the college objective, rather than because diversity increases quality. Research focused on identifying what college objective best fits the data is of much interest.[28] In this vein, it is interesting to compare the tuition results to the case where colleges maximize profits. It is not hard to show that a profit-maximizing college would have

[28] An example is whether need-based aid is better explained by colleges' desire for income diversity of their student body, as assumed in Epple, Romano, and Sieg (2006), or by third-degree price discrimination as hypothesized here. See the next point in the text.

a tuition function that is of the exact *form* of (14)! The main objective of the article is to determine the empirical content of the pricing equation in (14). Our estimation approach, discussed in detail in the next section, is therefore actually consistent with quality or profit maximization assumptions.

Distinguishing quality and profit maximization empirically would require distinguishing relatively subtle differences between equilibria under the two alternatives.[29] Given educational inputs, the quality maximizing college sets tuition to maximize profits, taking account of the peer value effects, so as to have the maximum funds to increase quality. However, the quality maximizing college has stronger incentive to spend on educational inputs, implying inputs will be higher in (14). Moreover, if the quality function is such that q_θ/q_I increases with I, as we might expect, then the weight on the ability-based peer effect $(\theta - b)$ in (15) will differ, implying the quality maximizer has stronger incentives to attract higher ability students. Likewise, the quality maximizer has stronger incentives to attract minorities.[30]

To test the implications of equation (14), we need to close the demand model and derive the conditional market shares for each private college. For that we need to derive the admission policies of state schools. From an empirical perspective, we require that public colleges adopt minimum ability admission thresholds that depend on the state of residence and the minority status of the student. Next we summarize a model of state colleges that generates state college policies that have these properties (with details in an appendix from the authors upon request).

From the perspective of a state college, a student with characteristics (m, b, y) is either an in-state student or an out-of-state student. We assume that tuition charged to in-state students is fixed exogenously at T_s and to out-of-state students at T_{so}, e.g., by the state legislature. Epple, Romano, Sarpca, and Sieg (2017) then show that state schools that maximize aggregate achievement of their in-state students use ability

[29] See Epple and Romano (1998, 2008) for an analysis of profit maximization by (secondary) schools in a related model.

[30] It is of interest to think about analogues to these results in other markets where for-profit and nonprofit providers coexist and might be more comparable, with hospitals a potential example.

threshold rules to determine access to the school. These thresholds differ for in-state and out-of state students. With $T_{so} > T_s$, as hold empirically, the ability threshold for admitting in-state versus out-of-state students can be higher or lower theoretically. Out-of-state are valued by state schools for their peer spillovers and higher regulated tuition. This can lead their admission threshold to be higher or lower depending on the relative weight in the quality function on the peer effect to the resource effect. Note, too, these admission thresholds will each be lower for minority status. In the empirical implementation, consistent with the data, we allow multiple state colleges that vary by quality and then let tuitions and admission thresholds vary among them. With these results in hand, we are now in a position to turn to empirical analysis and determine whether the predictions of this model regarding pricing and demand are consistent with the observed data.

4 Estimation

To estimate the model we need to invoke some additional parametric assumptions.

Assumption 1

a) The quality function is given by

$$q_j = \theta_j^{\gamma} \, I_j^{\omega} \, \Gamma_j^{\kappa} \, e^{u_j}, \quad \gamma, \omega, \kappa \; > \; 0 \tag{18}$$

where u_j is an unobserved exogenous characteristic.

b) The utility function is given by:

$$U_j(y - p_{sj} - L + A_{sj}, a(q_j, b)) = \alpha \ln(y - p_{sj} - L + A_{sj}) + \alpha \ln(q_j b^{\beta}) + \epsilon_j, \; \beta, \alpha > 0 \tag{19}$$

where α parameterizes the weight on the systematic component of utility.

c) The disturbances ε_j are independent and identically distributed with Type I Extreme Value Distribution.

The assumptions above then imply that the conditional choice probability for type (s, m, b, y) is given by, for j $\in J_a(m, s, b)$:

$$r_{sj}(m, b, y) \; = \; \frac{[(y - p_{sj}(m, b, y) - L + A_{sj}(y)) \; q_j]^{\alpha}}{\sum_{k \in J_a(m,s,b)} [(y - p_{sk}(m, b, y) - L + A_{sk}(y)) \; q_k]^{\alpha}}. \tag{20}$$

The pricing equation (12) for private colleges can then be written:

$$p_{sj}(m, b, y) \; = \; \frac{(1 - r_{sj})\alpha}{1 + (1 - r_{sj})\alpha} EMC_j(m, b) \; + \; \frac{1}{1 + (1 - r_{js})\alpha} \; (y - L + A_{sj}(y)) \tag{21}$$

Effective marginal costs at private colleges are given by:

$$EMC_j(m, b) = V_j' \; + \; I_j \; + \; \frac{\gamma I_j}{\omega \theta_j} \; (\theta_j - b) \; + \; \frac{\kappa I_j}{\omega \Gamma_j} (\Gamma_j^m - m) \tag{22}$$

The pricing function is then:

$$\begin{aligned} p_{sj}(m, b, y) \;\; &= \;\; \frac{(1 - r_{sj})\alpha}{1 + (1 - r_{sj})\alpha} \left(V_j' \; + \; I_j \; + \; \frac{\gamma I_j}{\omega \theta_j} \; (\theta_j \; - \; b) + \; \frac{\kappa I_j}{\omega \Gamma_j} (\Gamma_j - m) \right) \\ &+ \;\; \frac{1}{1 + (1 - r_{js})\alpha} \; \Big(y - L + A_{sj}(y) \Big) \end{aligned} \tag{23}$$

In addition we simplify notation by writing the marginal resource costs as

$$V_j \;\; = \;\; V'(k_j) \; + \; I_j \tag{24}$$

We treat the $V_1, ...V_J$ as additional parameters to be estimated.

The model implies an appealing decomposition of tuition. From (21), observe that tuition to student (s, m, b, y) is a convex combination of the student's effective marginal cost and cost adjusted income. The linkage of tuition to ability and minority status reflects the student's within college externality, which is independent of market power. The absence of market power would imply tuition equal to EMC, which arises only in the limit as $\alpha \to \infty$, implying idiosyncratic preferences are irrelevant.[31] The

[31] Note that tuitions would rise with quality absent market power because EMC increases with

linkage to income reflects market power over the student type (s, m, b, y). The weight on income increases with the student type's market share at the college indicating increased market power over the student. The weight on income decreases with α, the weight on the systematic component of utility. This indicates that market power declines as idiosyncratic preferences become less important.

The information set of the econometrician can be characterized as follows.

Assumption 2

- *We observe a sample $i = 1, ..., N$. Let s_i denote the state of student i, m_i the minority status, b_i ability, y_i income and p_{sji}, the tuition at college j. Note that we observe the tuition at the college attended in equilibrium, but not at colleges that are not attended in equilibrium. Let d_{ji} denote an indicator which is equal to one if student i attends college j and zero otherwise.*

- *L is known.*

- *θ_j, I_j, k_j are known for all j.*

- *In- and out-of-state tuitions at state colleges ($\bar{p}_{js}$) and posted prices at private colleges ($\bar{p}_j$) are known.*

- *$A_{sij}(y_i)$ are observed for all i and j.*

- *Prices for all students i at private colleges that are not paying the posted price are measured with classical error:*

$$\tilde{p}_{sji} = p_{sij}(m_i, b_i, y_i) + v_{ij} \tag{25}$$

where v_{ij} is iid across i and j.

I_j, θ_j, and Γ_j^m to the "average student," defined by $b = \theta_j$ and $m = \Gamma_j^m$.

Consider the subsample of students that attend private colleges and are not paying the full posted price, i.e. the subsample of students at private schools that obtain some institutional aid. Using this subsample we can identify and estimate most of the parameters of the model using the predictions of the model about price discrimination. In particular, we can implement the following semi-parametric estimator.

We estimate the conditional market shares $r_{sj}(m, b, y)$ for all students for the private college that is attended in the data. We use a simple flexible Logit estimator using a quadratic approximation in b and y, where the coefficients depend on m and s. We then use the estimated Logit model to predict the conditional choice probability denoted by $\hat{r}_{sj}(m, b, y)$. Alternatively, we could use non-parametric techniques such as kernel or sieve estimators.

Substituting the estimator of the conditional market share into the pricing equation, we obtain:

$$\begin{aligned} \tilde{p}_{sji} &= \frac{(1-\hat{r}_{sj})\alpha}{1+(1-\hat{r}_{sj})\alpha}\left(V_j + I_j + \frac{\gamma I_j}{\omega\theta_j}(\theta_j - b) + \frac{\kappa I_j}{\omega\Gamma_j}(\Gamma_j - m)\right) \\ &+ \frac{1}{1+(1-\hat{r}_{js})\alpha}\Big(y - L + A_{sj}(y)\Big) + v_{ji} \end{aligned} \tag{26}$$

where v_{ji} is the measurement error term. We can, therefore, identify and estimate α, the ratios γ/ω and κ/ω, as well as the marginal costs $V_1, ..., V_J$ using a semi-parametric NLLS estimator based on equation (26). We use a bootstrap algorithm to estimate the standard errors to account for the sequential nature of the estimation procedure.

All the empirical results reported on this article are based on this estimator. One nice property of this estimator is that it is consistent for large N, but small J. This scenario is relevant for most practical applications.

We do not allow for random coefficients in our utility specification because of data limitations.[32] It is difficult to identify the parameters of these distributions in a single cross-section without having access to second choice data as discussed in

[32]Epple, Jha, and Sieg (2018) exploit variations in the choice set due to school closures to estimate a demand model for secondary schools.

detail in Berry, Levinsohn, and Pakes (2004). We also do not observe large changes in the choice set over time or the full portfolio of colleges that a students chooses at the application stage. In principle, it is not difficult to extend our semi-parametric estimator to account for random effects in the demand side of the model.

For certain applications, knowledge of the level of ω, γ, and κ is useful. This section discusses how to identify and estimate the levels of these parameters using a modified version of the estimator suggested by Berry (1994).

Two additional challenges to estimation are present that are typically not encountered in standard demand analysis. First, the potential choice set of a student is unobserved by the econometrician. Our model implies, however, that both private and public schools use minimum-ability threshold rules to determine admission functions. These arise because both private and public colleges face binding price maxima that lead to minimum ability admission thresholds. We observe attendance in equilibrium and as a consequence can estimate minimum admission thresholds using order statistics. This allows us to characterize the relevant choice set for each student in the sample.

Second, private colleges engage in third-degree price discrimination. Hence institutional aid and net tuition policies of all private colleges are functions of income and ability as long as the price maximum is not binding. A key challenge encountered in estimation is that the institutional aid is observed only at the college that is attended in equilibrium. The econometrician does not observe the financial aid packages and, hence the net tuition, that were offered by the colleges that also admitted the student, but were ultimately rejected by the student. As a consequence, we cannot directly evaluate the conditional choice probabilities for each student. However, we can consistently estimate the institutional aid functions of each college type using nonparametric techniques such as kernel or sieve estimators. Given these consistent estimators we then can compute the conditional choice probabilities of each student.

Consider the full sample of all students, including those students that attend private colleges and that pay the posted price as well as students attending public colleges and universities.

We can construct the minimum ability threshold for each college, by computing the minimum ability of the students. Let our estimator be denoted by b_{jm}^{min}. We can then identify the choice set for all students as follows:

$$J_a(m, s, b) = \{s | b \geq b_{sm}^{min}\} \cup \{o \in S \setminus \{s\} | b \geq b_{om}^{min}\} \cup \{j \in P | b \geq b_{jm}^{min}\} \cup \{0\} \quad (27)$$

The first and second sets are, respectively, the in-state public colleges and the out-of state public colleges admitting the student. The third set denotes the set of all private colleges to which the student is admitted, and the last set is the outside option.

We then non-parametrically estimate the prices for each student at each college to which the student was admitted based on the observed tuition levels, using a local quadratic polynomial. Let us denote these estimates by $\hat{p}_{sji}^{np}$.

For private colleges, we use a local polynomial smoothing estimator to estimate the tuition function. The polynomial constructs a non-parametric estimate of tuition based on ability and income, and interpolates only where the observed data span. That is, if we observe an individual with a similar ability and income attending the college with a given tuition, the LOESS estimator calculates a polynomial relationship among tuition, ability, and income within the relevant bandwidth and predicts tuition locally. However, the resulting admission set–where a tuition can be predicted–is really a combination of admission and matriculation. Thus, many lower quality colleges would appear to reject high-quality applicants, because no such applicants are observed at the college. Hence we assume that if a college accepts an individual with ability b^{min}, it accepts all individuals where $b_i \geq b^{min}$. Then, we extrapolate the local polynomial to such individuals to ensure all admitted individuals have a valid first-stage tuition offer.

Substituting the nonparametric estimates of the tuitions into the conditional choice probabilities, we obtain

$$\hat{r}_{ji} \ = \ \frac{[(y_i - \hat{p}^{np}_{sji} - L + A_{s_i,j}(y_i))q_j]^\alpha}{\sum_{k\in J_a(m_i,s_i,b_i)}[(y_i - \hat{p}^{np}_{ski} - L + A_{sk}(y_i))q_k]^\alpha} \tag{28}$$

The quality levels for each school are determined by the fixed point of the following mapping:

$$\tilde{q}_j = q_j + ln(s^N_j) - ln(s_j(q)) \quad j = 1, ..., J-1 \tag{29}$$

where: q_j is initial guess of the quality, s^N_j is the average empirical market share of college j observed in the data, and $s_j(q)$ is the predicted average market share using the initial guess about the vector of qualities:

$$s_j(q) = \frac{1}{N}\sum_{i=1}^{n} \hat{r}_{ji} \tag{30}$$

We can identify q_j's for each college, subject to a normalization such as $q_1 = 1$. The normalization of quality is necessary as market shares add up to one.

Using the fact that $q_j = \theta^\gamma_j \ I^\omega_j \Gamma^\kappa_j e^{u_j}$ we obtain the the following regression model:

$$\ln(q_j/q_1) = \omega\left(\frac{\gamma}{\omega}\ln(\theta_j) + \frac{\kappa}{\omega}\ln(\Gamma_j) + \ln(I_j) - \frac{\gamma}{\omega}\ln(\theta_1) - \frac{\kappa}{\omega}\ln(\Gamma_1) - \ln(I_1)\right) + u_j - u_1 \tag{31}$$

Define

$$w_j = \frac{\gamma}{\omega}\ln(\theta_j) + \frac{\kappa}{\omega}\ln(\Gamma_j) + \ln(I_j) - \frac{\gamma}{\omega}\ln(\theta_1) - \frac{\kappa}{\omega}\ln(\Gamma_1) - \ln(I_1) \tag{32}$$

and note that w_j is known at this point. Rewriting equation (31) as

$$\ln(q_j/q_1) = \omega \ w_j \ + \ u_j - u_1 \tag{33}$$

and hence ω can be estimated using a simple ratio estimator. This estimator is consistent despite the fact that w_j and $u_j - u_1$ may not be independent because the regression above does not have an intercept. Note that the last step of the estimator requires a large number of colleges or preferably multiple markets.

Finally, consider identification of college cost functions. Note that the fixed costs of operating a private college F are not identified from our analysis. One could impose a functional form assumption on the marginal custodial costs such as $V'(k) = c_0 + c_1 k_j + \omega_j$, where k_j is the size of college j. One could then estimate the parameters c_0 and c_1 using the following equation:

$$V_j - I_j \;=\; c_0 + c_1 k_j + \omega_j \tag{34}$$

With a large enough sample of private school clusters and a suitable instrument for k_j, one could then identify and estimate c_0 and c_1. (Note that we identify V_j from the price functions and I_j is observed). Epple, Romano and Sieg (2006) use endowment income as an instrument for k_j. The cost function for public schools is not identified from our approach.

5 Parameter Estimates

Table 4 summarizes the parameter estimates for our model. Note that these estimates are based on the subsample of students at private universities that received a positive amount of institutional financial aid. The relevant sample size is 2,270. We only need the other students to construct the correct measures of the conditional market shares, which then enter into the pricing equation. We estimate three types of model specifications. The first two specifications ignore minority status whereas the third specification – which is our preferred specification – also accounts for minority status. Dollar amounts are measured in tens of thousands.

The first column uses unweighted data. The second and third column uses the weights suggested by NPSAS. Comparing the weighted estimates in Column 2 with the unweighted estimates in Column 1, we find only small differences in the estimated parameter values. The main difference is that the unweighted estimator yields a somewhat greater point estimate of α. The estimates in Column 2 are similar to the ones in Column 3.

Focusing on our preferred estimates in Column 3 we obtain an estimate of α which is equal to 72.72 with an estimated standard error of 7.13. As a consequence we find that our estimate is highly significant at standard levels of significance. Note that α is primarily identified from the observed pricing by income.

Next consider the ratio of $\frac{\gamma}{\omega}$. Our point estimate equals 0.079 with an estimated standard error of 0.012. This ratio is primarily identified off the observed merit based aid. We conclude that both key parameters are estimated with high precision. The quantitative implications of these estimates are discussed below.

Our point estimate for the ratio of $\frac{\kappa}{\omega}$ is 0.012 with a standard error of 0.003. This ratio is primarily identified off the observed institutional aid to minority students holding income and ability fixed. The average predicted marginal effect of being a minority student in our model is a \$900 discount. We conclude that our model provides strong evidence that private schools care about racial diversity.

We can also estimate the marginal resource costs of admitting an additional student to the college. Not surprisingly, we find that there is much heterogeneity in marginal costs. Our estimates range between approximately \$5,400 and \$16,600. Note that these estimates combine marginal expenditures on educational inputs and marginal custodial costs.

When we implement the last two stages of our estimator, we obtain point estimates for ω that range between 0.027 and 0.033. But recall that our sample size is small ($J = 15$). Hence we need to interpret these estimates with caution.

We conclude that our estimator performs well and provides reasonable estimates for the key parameters. Next we explore the implications of our model estimates for price discrimination and market power in the market of higher education.

6 Price Discrimination in U.S. Higher Education

Table 5 reports the average derivatives of price in ability and in income implied by our model and compares these predictions to simple reduced form OLS estimates with fixed cluster effects. It is instructive to compare these estimates, which embody the structural properties of our model, to OLS estimates that simply assume linear pricing. Overall, we find that our model captures well the basic regularities in the data revealed by OLS. The payoff from our model is that we can assess the role of local market power and nonlinearities in the patterns of pricing by income and ability that are observed in the data.

To gain some additional insights into the predicted magnitude of pricing by income and ability as well as the extent of market power, it useful to decompose the prices paid by students into the different components. For students not paying the posted price, the marginal effect of ability on price is approximately given by:

$$\frac{\partial p_{sj}(m,b,y)}{\partial b} \approx -\frac{(1-r_{sj})\alpha}{1+(1-r_{sj})\alpha}\,\frac{\gamma I_j}{\omega\theta_j} \tag{35}$$

The marginal effect of income on price is approximately:

$$\frac{\partial p_{sj}(m,b,y)}{\partial y} \approx \frac{1}{1+(1-r_{js})\alpha} \tag{36}$$

We compute the "mark-up" as the difference between price and effective marginal cost:

$$\text{mark-up}_j(s,m,b,y) = p_{sj}(m,b,y) \; - \; EMC_j(m,b). \tag{37}$$

Computing the mark-up over EMC, rather than the more standard mark-up over marginal resource cost, is to capture market power by purging the discounts to ability and minority status that reflect within college externalities.

The first row of Table 6 shows the value of the average within cluster mark-up, i.e., the average within cluster value of (35) implied by our estimates and observed college attendance. Similarly calculated, using (33) and (34) respectively, the second and third rows report within cluster averages of marginal ability and income pricing effects. Looking across the columns, average mark-ups rise rapidly along the quality hierarchy, ranging above \$13,000 for elite colleges. Our estimates imply little difference in average of marginal pricing by income. These average effects range between \$210 and \$510 for an increase of \$10,000 in family income among the 10 clusters. However, our estimates imply much more variation in pricing by ability and in the mark-ups. Cluster averages of marginal pricing by ability ranges between tuition reductions of \$920 to \$1,960 (for an increase in one standard deviation in ability). As reported in the last row, the largest discounts for minority status occur at the four highest average-ability schools, with discounts ranging from \$1,600 to \$5,750. As these discounts stem mainly from differences in EMC, one can see from (24) that larger minority discounts result from a combination of having fewer minority students (lower Γ_j) and from higher I_j. Overall average mark-ups range between 3.5% and 35.5%. The large mark-ups allow us to conclude that the most selective colleges have significant market power.

The average estimates mask price discrimination within colleges, which is more substantial. To illustrate the magnitude of these effects, we focus on Clusters 1 and 2, which include the most selective colleges. Table 7 reports "local" market shares for non-minority students at the two most selective private colleges, with these market shares conditional on deciles for income and ability.[33] Because colleges value student ability and price discriminate according to income and ability, the equilibrium exercise of market power will vary with student characteristics. The conditional market shares for high ability and high income students in these clusters are much larger than the overall unconditional market share, which is equal to 0.08. As a consequence the college has significantly larger local market power than is suggested by its overall market share.

[33] We use deciles, though localness can be defined using a finer or broader delineation.

The panels of Table 8 report predicted mark-ups by income and ability for majority students in Clusters 1 and 2. [34] The values in the upper panel are the predicted mark-ups over EMC for a student at the midpoint of the cell, averaged over clusters 1 and 2, the cells delineated by quintiles in these colleges calculated using our estimated admission thresholds. [35]

We see that these mark-ups increase precipitously with income, especially as students become relatively wealthy. A combination of higher willingness to pay as income rises and increasing market power of top tier colleges over richer students explains the latter. The mark-ups to the richest students of relatively low ability are very high, about $18,000. For given income, mark-ups vary little with ability, except for the richest quintile where they fall substantially as ability increases (see the rightmost column). This reflects more competition for and thus less market power over high ability students that have enough income to be willing to pay for expensive colleges.

The lower panel calculates the more traditional mark-ups over marginal resource costs. That tuition is higher for low ability students and lower for high ability students is not purged in this version of the mark-up, and these mark-ups are then higher (lower) for low (high) ability students than in the upper panel. For high ability students without much income, these mark ups are negative, implying fellowships. Overall, we see that mark-ups and market power vary widely with student type.

We thus conclude that the most selective colleges have significant market power and charge significant mark-ups to students, especially for higher income, majority students. Is the exercise of market power by private colleges inefficient given the assumption of quality maximization? We show in an appendix (available on request) that tuition equal to effective marginal cost is first-best efficient. Thus, a case for

[34] The entries left blank in the table are because the number of students that actually attend within the cell, and thus used in estimation, are so low that the predictions are not credible.

[35] We use the estimated admission thresholds to calculate the minimum ability, the bottom value of the lowest ability quintile. Then we use the data for students above this minimum ability attending all colleges to compute ability quintile thresholds and income quintile thresholds, thus delineating the cells.

social concern about college pricing exists. But several caveats are noteworthy. For one, quality maximizers spend their "profits" on educational resources, which benefits all their students. Moreover, higher income students pay more. Given that educational resources in a college are evenly distributed, an implicit transfer from richer to poorer students arises. This is, of course, not the ideal way to enact such transfers assuming they are socially desirable. [36] A separate issue is that social externalities from educational achievement, not an element of our model, would imply higher educational expenditure than may be feasible with tuitions equal to effective marginal cost. Further assessment of the efficiency question would require more modeling and assumptions and might imply difficult empirical challenges. This is an interesting topic for further research.

7 Policy Analysis

Markups are a function of the admissions offers faced by students and hence depend on state policies towards higher education. Recall that our empirical model allows each state to have up to four different types of public universities that are differentiated by tuition and quality.[37] Some states invest heavily into the quality of their public universities whereas others rely mostly on private and out-of-state public universities to educate their students. In this section, we determine how the quality and pricing of local four-year public colleges impacts students' decisions and welfare.

To conduct this policy analysis, we simulated a sample from the distribution of students in a state with only the lowest quality of public school in-state. The simulation exercise also requires us to draw idiosyncratic shocks for each element in the choice set. Note that the students in our sample can attend public schools of any quality by traveling to another state, but face differing prices and admissions thresholds by

[36] The $p = EMC$ result for first-best efficiency assumes lump-sum income transfers to maximize the social welfare function.

[37] See appendix B for details about the availability of public colleges and universities.

in-state status. If those students were transplanted to a state where all four qualities of public school are available in-state, they face lower prices and more lenient admissions thresholds. This has a number of potential effects: students can continue at their same public or private school, potentially facing lower prices, or they can switch schools – to a better quality public school, to a public school instead of a private school, or to a public school instead of community college.

Table 9 summarizes our findings of this policy analysis. It reports changes in tuition and peer quality as well as the compensating variation associated with this policy change. Note that all students are weakly better off after the policy change.

We consider three groups of students. The first group consists of students that initially attended community colleges. Given only a weak public university in the state in question, it has a large population attending community college. Many of these students would continue to attend community college after switching states, about 73% of the total students in the state. However, 35% of these students choose to move from community college to public school. These students face higher prices (an average tuition increase of $3,500). They also experience a nearly 0.50 standard deviation increase in peer quality and a modest welfare gain. The second (small) group consists of students that initially attended public universities. Most of these students switch to a better quality public school, increasing their tuition by about $5,000 in exchange for a 0.30 standard deviation increase in peer quality. They experience a large welfare gain, equal to $2,346 as measured using CV.

Of the remaining students, 63% (17% overall) attend the same private school, but face slightly lower prices due to decreased market power, an average savings of $400 per student. Another 2% of students switch to public schools from private school. These students experience a decline in average peer quality. They also save an average of $11,000 in tuition, with the largest welfare gain of $6,828. Overall, the average gain from providing access to a more differentiated system of public education is significant, approximately $465 per student who benefits. The average gain overall is

$127.

Table 10 summarizes the distribution of these effects. Reported are within cell averages of tuition changes and peer group changes for all students, including those that do not switch school type. High ability students are the most likely to experience tuition savings, whereas those at the lowest end of the ability distribution generally experience an increase in tuition (in exchange for higher peer quality). The highest ability students also experience a gain in average peer quality, whereas the students in the middle of the state ability distribution generally do not switch schools.

8 Conclusions

We have shown how to extend standard concepts of mark-ups and market power to model with non-for-profit organizations that maximize quality instead of profits. We have developed a new semi-parametric estimator for a model that explains price discrimination and market power in the U.S. market for higher education. We have implemented our new estimator using data from the NPSAS. We obtain reasonable estimates for all of the key parameters. Our empirical findings suggest that the majority of private colleges in the U.S. engage in pricing by income, ability, and minority status. A $10,000 increase in family income increases tuition at private schools by an average of $210 to $510. A one standard deviation increase in ability decreases tuition by approximately $920 to $1,960 depending on the selectivity of the college. There are large and substantial discounts for minority students which range between approximately $110 (at historically black colleges) and $5,750 dollars. Average mark-ups in colleges range between 3.5 and 33.5 percent, but vary substantially within colleges and are very large for high income students. There is much more heterogeneity in mark-ups within colleges than among colleges. Our analysis suggests that highly selective colleges have significant market power, especially for high income, high ability, non-minority students. Finally, we have conducted some policy experiments and shown that there are significant gains from having access to a diversified set of public schools. There are direct effects for students attending these schools. In addition, there are indirect effects for students that chose to attend private schools in equilibrium as they benefit from the fact that they have better outside options.

We view the results of this article as promising for future research. One might add private and state colleges with alternative objectives, perhaps focused on serving a segment of the population. Making endogenous state college pricing is of interest, a challenge being modeling public policy choice. Consideration of foreign students is of interest. In practice, there are small differences among the colleges within a cluster. Colleges may engage in some limited competition for students that our model fails to capture. As a consequence, our empirical analysis may understate the level of competition in the market for higher education.

References

Arcidiacono, P. (2005). Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?. *Econometrica*, *73 (5)*, 1477–1524.

Avery, C. and Levin, J. (2010). Early Admissions at Selective Colleges. *American Economic Review*, *100 (6)*, 1–34.

Bajari, P., Hong, H., and Nekipelov, J. K. D. (2010). Estimating static models of strategic interactions. *Journal of Business and Economic Statistics*, *28(4)*, 469–482.

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal of Economics*, *25(2)*, 242–262.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63 (4)*, 841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market. *Journal of Political Economy*, *112*, 68–104.

Bettinger, E., Long, B. T., Oreopolous, P., and Sanbonmatsu, L. (2005). Family Incomes and Net Prices at Highly Selective Private Colleges and Universities. *Journal of Human Resources*, *60 (4)*, 769–790.

Bordon, P. and Fu, C. (2015). College-Major-Choice to College-Then-Major-Choice. *Review of Economic Studies*, *82*, 1247–88.

Cestau, D., Epple, D., and Sieg, H. (2017). Admitting Students to Selective Education Programs: Merit, Profiling, and Affirmative Action. *Journal of Political Economy*, *125 (3)*, 761–797.

Chade, H., Lewis, G., and Smith, L. (2014). Student Portfolios and the College Admissions Problem. *Review of Economic Studies*, *81*, 971–1002.

Dynarski, S. (2002). The behavioral and distributional implications of aid for college. *American Economic Review*, *92 (2)*, 279–85.

Ehrenberg, R. (2002). *Tuition Rising*. Harvard University Press.

Epple, D., Jha, A., and Sieg, H. (2018). The Superintendent's Dilemma: Managing School District Capacity as Parents Vote with Their Feet. *Quantitative Economics*, *9 (1)*, 483–520.

Epple, D. and Romano, R. (1998). Competition Between Private and Public Schools, Vouchers and Peer Group Effects. *American Economic Review*, *88*, 33–63.

Epple, D. and Romano, R. (2008). Educational Vouchers and Cream Skimming. *International Economic Review*, *49*, 1395–1435.

Epple, D. and Romano, R. (2010). Peer Effects in Education: a Survey of the Theory and Evidence. In Benhabib, J., Bisin, A., and Jackson, M. (Eds.), *Handbook of Social Sciences, Vol 1b*, pp. 1053–1163. Elsevier Science.

Epple, D., Romano, R., Sarpca, S., and Sieg, H. (2017). A General Equilibrium Analysis of Access to Higher Education in the U.S.. *Journal of Public Economics*, *55*, 164–178.

Epple, D., Romano, R., and Sieg, H. (2003). Peer Effects, Financial Aid, and Selection of Students into Colleges. *Journal of Applied Econometrics*, *18 (5)*, 501–525.

Epple, D., Romano, R., and Sieg, H. (2006). Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. *Econometrica*, *74(4)*, 885–928.

Fillmore, I. (2015). Price discrimination and public policy in the U.S. college market. Working Paper.

Fu, C. (2014). Equilibrium Tuition, Applications, Admissions and Enrollment in the College Market. *Journal of Political Economy*, *122 (2)*, 225–81.

Goutte, C., Toftand, P., Rostrup, E., Nielsen, F., and Hansen, T. (1999). On clustering fMRI time series. *NeuroImage*, *9 (3)*, 298–310.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, *47 (1)*, 153–161.

Hickman, B. (2013). Pre-college Human Capital Investment and Affirmative Action. Working Paper.

Hoxby, C. and Turner, S. (2015). What High-Achieving Low-Income Students Know About College Options?. *American Economic Review: Papers and Proceedings*, *105 (5)*, 514–17.

Kapor, A. (2016). Distributional Effects of Race-Blind Affirmative Action. Working Paper.

MacLeod, B. and Urquiola, M. (2015). Reputation and school competition. *American Economic Review*, *105(11)*, 3471–3488.

Mardia, K., Kent, J., and Bibby, J. (1979). *Multivariate Analysis*. Academic Press.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*. Academic Press.

McPherson, M. and Schapiro, M. O. (1998). *The Student Aid Game*. Princeton University Press.

McPherson, M. and Schapiro, M. O. (2006). U.S. Higher Education Finance. In *Handbook of Education*. Elsevier North Holland.

Rothschild, M. and White, L. (1995). The Analytics of the Pricing of Higher Education and Other Services in Which the Customers are Inputs. *Journal of Political Economy*, *103*, 573–623.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far?. In Hanushek, E. e. a. (Ed.), *Handbook of Education 3*, pp. 249–277. Elsevier Science.

Sieg, H. and Wang, Y. (2018). The Impact of Student Debt on Education, Career, and Marriage Choices of Female Lawyers. *European Economic Review*, *forthcoming*.

van der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression Discontinuity Approach. *International Economic Review*, *43 (4)*, 1249–87.

A Construction of the Ability Measure and Ability Thresholds

We measure ability by predicting students' first-semester GPA as a function of their high school GPA, ACT score (or SAT score converted to ACT score), gender, major, and college choice.

The predicted GPA for student i at school j in discipline d is given by:

$$GPA_{ijd} = \beta_0 + \beta_1 \text{HSGPA}_i + \beta_2 \text{ACT}_i + \beta_3 \text{HSGPA}_i \times \text{ACT}_i + \beta_4 \text{female}_i + \beta_j + \beta_d + \epsilon_{ijd} \tag{38}$$

where β_j represents a college fixed effect and β_d represents a major fixed effect (12 majors, humanities omitted). Using a sample of approximately 5,000 white students at 4-year public or private universities, we obtain the following prediction (for a generic discipline at a generic school as these fixed effects and the intercept are dropped):

$$\widehat{GPA_i} = -3.184 \text{HSGPA}_i + -2.559 \text{ACT}_i + 0.918 \text{HSGPA}_i \times \text{ACT}_i + 21.961 \text{female}_i \tag{39}$$

The R^2 for the estimated model (including fixed effects) is 0.9342. After clustering, we then standardize the ability measure to have standard deviation 1 and mean 0.415, such that all schools have $\theta_j \geq 0$.

There are no explicit ability admission thresholds in the data. We estimate these thresholds using all students except a small number of "legacy students" at some of the most selective universities.[38] We construct these by taking the first percentile predicted GPA at any public college with at least ten non-minority students (as our model implies different admission thresholds for minority students), and applying

[38] Legacy students are believed to contribute additional value to the school (perhaps through alumni donations), and thus are subject to different admission criteria. Due to this unobserved characteristic, they have a lower net marginal cost for the school and may be admitted despite lower ability. Legacy students are identified off the empirical CDF of ability within a school–they precede the lowest flat region in the CDF. In support of our hypothesis, such students tend to be non-minority students with high family income.

the minimum within cluster as the cluster admission threshold. Any students below this threshold presumably have characteristics desirable to a college, and so these students are "bumped up" to the threshold predicted GPA. This new measure is used in admission and tuition estimation for all colleges. We considered other approaches to construct these admission thresholds, and found that the main results reported in this article are not sensitive to the specifics discussed above.

B Public School Availability

Public colleges sampled by NPSAS do not comprise the universe of public colleges in the U.S., whereas the availability of state colleges for a student affects choice probabilities and also pricing by private colleges in the model. Therefore, we determine the availability of each state college type within each state by adding in the schools from the top 100 public schools in the U.S. (via U.S. News and World Report) that were not included in the NPSAS. These schools were allocated to the appropriate cluster by distance from the cluster center. The table below summarizes state college availability in each state in the NPSAS and our additions (printed in red) by cluster.

Tables

Table 1: Sample Selection

	2-year public	4-year public	4-year private	total
Full NPSAS 2012	31,000	17,300	9,010	57,300
First-year only	17,860	4,530	4,210	26,590
No atypical attendance	5,380	3,370	3,470	12,220
No athletes	5,330	3,310	3,280	11,910
No veterans	5,190	3,230	3,230	11,660
No missing ability	4,180	3,160	3,170	10,510
No missing state	4,150	3,130	3,090	10,370
No missing school expenditures	3,510	2,910	3,070	9,490

Note: Unweighted counts rounded to nearest 10 as per NCES policy.

Table 2: Selected Characteristics for NPSAS 2012 Sample

	2 year public	4 year public	4 year private	all
Number of students	3,510	2,910	3,070	9,490
Number of students *(weighted)**	521,638	583,844	342,519	1,448,001
Number of Colleges	300	250	350	900
Number of Colleges *(weighted)***	1549	713	1286	3548
Average ACT Score	19.72	21.88	23.79	21.55
Average Ability	0.00	0.45	0.81	0.37
Average In-state Net Tuition***	3.00	5.73	26.37	12.02
Average Out-of-state Net Tuition	6.48	15.48	26.37	15.50
Average Income	48.4	76.9	94.8	70.9
Female	0.53	0.54	0.57	0.55
Black	0.18	0.17	0.14	0.17
Hispanic	0.19	0.13	0.11	0.15

*Students are weighted to be nationally representative, using inverse probability weights provided by the NCES. All other student-level statistics (e.g. ACT score, gender) are also weighted.
**Colleges are weighted to be nationally representative, using inverse probability weights provided by the NCES. Tuition values are also weighted.
***Tuition and income reported in $1,000s.

Note: Unweighted counts rounded to nearest 10 as per NCES policy.

Table 3: Characteristics of Clusters

cluster	in-state admit	out-state admit	mean ability	mean ACT	mean posted	mean tuition	instruct. expend.	percent black	percent Hispanic	count colleges	count students	weighted students
Private 4-Year Colleges												
1	-0.40	-0.40	1.66	28.59	39.31	25.28	37.96	0.07	0.11	20	450	36,758
2	0.35	0.35	1.48	27.77	41.63	29.75	17.30	0.06	0.10	20	290	38,264
3	-1.42	-1.42	0.93	24.81	30.74	19.30	12.86	0.03	0.09	10	130	16,269
4	-1.26	-1.26	0.82	24.47	36.66	22.25	11.52	0.08	0.11	40	420	45,429
5	-1.88	-1.88	0.76	23.07	23.76	15.41	9.07	0.16	0.11	40	330	30,431
6	-1.42	-1.42	0.61	22.61	31.11	17.26	8.34	0.16	0.16	50	390	51,837
7	-1.86	-1.86	0.49	21.80	26.73	14.47	6.66	0.14	0.09	60	490	49,517
8	-1.87	-1.87	0.43	21.33	18.22	12.07	6.29	0.18	0.10	30	170	27,424
9	-1.61	-1.61	0.39	21.09	21.78	11.57	5.42	0.19	0.12	40	240	26,491
10	-1.45	-1.45	0.22	20.93	12.19	8.18	5.47	0.36	0.06	30	170	20,099
Public 4-Year Colleges												
11	-0.53	-0.53	0.69	23.05	15.52	13.18	10.43	0.05	0.19	10	140	31,538
12	-1.73	-1.17	0.58	22.50	11.17	9.33	9.36	0.13	0.08	60	840	165,888
13	-1.76	-1.42	0.43	22.04	7.33	6.06	7.50	0.15	0.15	110	1,180	242,419
14	-2.38	-1.49	0.27	20.64	4.31	3.50	6.05	0.28	0.15	80	750	143,998
Public 2-Year Colleges												
15			0.00	19.72	3.18	2.98	4.48	0.18	0.19	300	3,510	521,638

Instructional expenditures weighted by institutional weight. All other means weighted by individual weight
Unweighted counts rounded to the nearest 10 as per NCES policy.
Tuition and expenditures reported in $1,000.

Table 4: Parameter Estimates

	(1)	(2)	(3)
Weights	No	Yes	Yes
Minority Status	No	No	Yes
α	86.56***	70.26***	72.72***
	(8.58)	(6.68)	(7.13)
$\frac{\gamma}{\omega}$	0.074***	0.0734***	0.079***
	(0.012)	(0.012)	(0.012)
$\frac{\kappa}{\omega}$			0.01***
			(0.003)
V_1	1.22***	1.21***	1.23***
	(0.07)	(0.07)	(0.07)
V_2	1.69***	1.65***	1.66***
	(0.07)	(0.07)	(0.07)
V_3	1.43***	1.40***	1.41***
	(0.08)	(0.08)	(0.08)
V_4	1.82***	1.81***	1.82***
	(0.05)	(0.05)	(0.05)
V_5	1.15***	1.14***	1.14***
	(0.05)	(0.05)	(0.05)
V_6	1.48***	1.46***	1.46***
	(0.04)	(0.04)	(0.04)
V_7	1.15***	1.13***	1.14***
	(0.04)	(0.04)	(0.04)
V_8	0.93***	0.92***	0.92***
	(0.07)	(0.07)	(0.07)
V_9	1.09***	1.08***	1.08***
	(0.05)	(0.05)	(0.05)
V_{10}	0.56***	0.54***	0.54***
	(0.08)	(0.08)	(0.08)
Note $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$			

Table 5: Pricing by Ability and Income

	(1)	(2)	(3)
Weights	No	Yes	Yes
Minority Status	No	No	Yes
Structural Estimates of Pricing by Ability and Income			
$\frac{\partial p}{\partial b}$	-0.095***	-0.105***	-0.112***
$\frac{\partial p}{\partial y}$	0.013***	0.015***	0.014***
Reduced Form Estimates of Pricing by Ability and Income			
$\frac{\partial p}{\partial b}$	-0.113***	-0.112***	-0.121***
$\frac{\partial p}{\partial y}$	0.017***	0.016***	0.016***
Note *p<0.1; **p<0.05; ***p<0.01			
OLS estimates account for cluster fixed effects			

Table 6: Predicted Mark-ups and Pricing by Income, Ability, and Minority Status

	(1)	(2)	(3)	(4)	(5)
markup	13.16	13.22	5.30	4.11	4.05
ability	-1.80	-0.92	-1.11	-1.12	-0.94
income	0.03	0.03	0.02	0.02	0.03
minority status	-5.75	-3.08	-4.23	-1.60	-0.58
	(6)	(7)	(8)	(9)	(10)
markup	2.66	3.09	2.86	0.75	2.77
ability	-1.06	-1.06	-1.14	-1.09	-1.96
income	0.05	0.04	0.04	0.05	0.03
minority status	-0.51	-0.50	-0.33	-0.27	-0.11

Note: Markups include pricing by minority status. Figures (in $1,000) calculated using full sample, not just those observed to receive aid.

Table 7: Local Market Shares in Clusters 1 and 2

	income percentile									
ability	10	20	30	40	50	60	70	80	90	100
10	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.00
20	0.01	0.02	0.00	0.00	0.00	0.03	0.01	0.02	0.03	0.06
30	0.01	0.00	0.00	0.00	0.02	0.05	0.00	0.04	0.03	0.05
40	0.03	0.02	0.01	0.01	0.04	0.01	0.01	0.00	0.04	0.02
50	0.02	0.01	0.00	0.02	0.02	0.06	0.03	0.04	0.09	0.11
60	0.00	0.02	0.01	0.00	0.02	0.04	0.08	0.05	0.04	0.05
70	0.00	0.01	0.01	0.02	0.06	0.05	0.03	0.10	0.08	0.08
80	0.04	0.08	0.05	0.07	0.07	0.11	0.04	0.05	0.07	0.24
90	0.08	0.07	0.04	0.07	0.13	0.20	0.10	0.07	0.20	0.23
100	0.22	0.28	0.37	0.18	0.29	0.28	0.38	0.31	0.41	0.53

Note: Table gives proportion of each income-ability percentile combination attending colleges in Cluster 1 or 2. Proportions are unweighted.

Table 8: Predicted Mark-ups by Ability and Income Quintile, Clusters 1 and 2

	Mark-ups over Effective Marginal Cost				
ability\income	0%-20%	20%-40%	40%-60%	60%-80%	80%-100%
0%-20%	-	0.22	0.63	1.15	17.99
20%-40%	-	0.23	0.64	1.16	17.56
40%-60%	-	0.24	0.65	1.17	15.56
60%-80%	-	0.25	0.66	1.18	12.66
80%-100%	0	0.26	0.68	1.20	9.15
	Mark-ups over Marginal Resource Cost				
ability\income	0%-20%	20%-40%	40%-60%	60%-80%	80%-100%
0%-20%	-	2.50	2.92	3.44	20.28
20%-40%	-	1.89	2.30	2.82	19.22
40%-60%	-	1.32	1.73	2.25	16.64
60%-80%	-	0.68	1.09	1.61	13.10
80%-100%	-0.92	-0.60	-0.19	0.33	8.28

Table 9: Switching induced by change of state

	Fraction	Δ Tuition	Δ Peer Quality	CV†
Initially private				
Same private	0.17	-400	0.00	400
Move to better private	None			
Move to public	0.01	-11,107	-0.30	6,828
Move to CC	None			
Initially public				
Same public	None			
Move to better public	0.00*	4,946	0.30	2,346
Move to private	None			
Move to CC	None			
Initially CC				
Same CC	0.73	0	0.00	0
Move to private	None			
Move to public	0.09	3,518	0.48	184
Average		222	0.04	465

Notes: Switching induced by moving students from a state with only the lowest quality of public school instate to a state with all public schools available instate. Tuition change is reported in dollars (negative implies savings); peer quality changes are reported in standard deviations.
† "CV" is the negative of compensating variation–the dollar-equivalent utility gain from switching schools. The CV column "Average" conditions on students whose welfare strictly improves.
*Less than 0.5% of students fall into this category.

Table 10: Distribution of switching gains by ability and income quintile

Panel A: Change in tuition

ability\income	1	2	3	4	5
1	496	665	589	758	339
2	552	1,007	1,017	905	919
3	-1	0.00	0	0.00	-5
4	-49	0.00	0	-1	-27
5	-805	-271	191	-237	-466

Panel B: Change in peer quality

ability\income	1	2	3	4	5
1	0.08	0.11	0.10	0.13	0.06
2	0.08	0.11	0.12	0.11	0.11
3	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00
5	0.00	0.03	0.06	0.01	0.01

Notes: "1" represents the lowest quintile; "5" represents the highest quintile. Switching induced by moving students from a state with only the lowest quality of public school instate to a state with all public schools available instate. Tuition change is reported in dollars (negative implies savings); peer quality changes are reported in standard deviations. Figures are the average for all students within the cell.

Table 11: Public College Availability*, **

	1	2	3	4	All	Non-NPSAS school
AK			1		1	
AL	4	1	2		7	
AR	1	3			4	
AZ	1		1		2	
CA	2	14	2	1	19	UC Santa Barbara, UC Riverside
CO	1	3	1	1	6	
CT		3	1		4	UCONN
DE			2		2	
FL		7	1		8	
GA	1	9	1		11	
HI			1		1	Univ HI-Manoa
IA	1		1		2	Univ IA
ID		4			4	
IL	2		5	1	8	U IL Urbana
IN	1	3	3		7	
KS	1	3	1		5	
KY	2	1	2		5	
LA	1	6			7	LSU
MA		5	1	2	8	Umass Amherst, Lowell
MD	1	4	1	1	7	
ME	1				1	
MI	9		1	1	11	
MN		5	1		6	Univ MN-Twin Cities
MO		3	1	1	5	MO S&T
MS	1	4			5	
MT		2			2	
NC		9	1	1	11	NC State-Raleigh
ND		1	1		2	
NE	1	3			4	
NH	1				1	
NJ	5		1	2	8	Rutgers
NM		2			2	
NV			1		1	
NY	2	9	3		14	SUNY Binghamton, Buffalo
OH	6	1	2		9	
OK		5	1		6	
OR	2		1		3	
PA	6	1	2		9	
RI	1				1	
SC	4	1	1		6	
SD		2			2	
TN	1	4	1		6	
TX	1	16	1	1	19	Texas A&M-College Station
UT		4	1		5	
VA	1	4	3	1	9	UVA
VT			1		1	
WA	4			1	5	
WI	4	2	1		7	UWI-Madison
WV	1	2			3	
WY			1		1	U WY
Grand Total	**70**	**146**	**53**	**14**	**283**	

* In ascending order of average ACT score, clusters are 2,1,3,4
**Red numbers mark an addition from US News & World Report top 100 public schools (clustered based on centroid distance). Schools not added if state already had at least one of that type (e.g. Berkeley in CA).