COLLEGE TUITION AND INCOME INEQUALITY*

Zhifeng Cai
Rutgers University

Jonathan Heathcote
FRB Minneapolis and CEPR

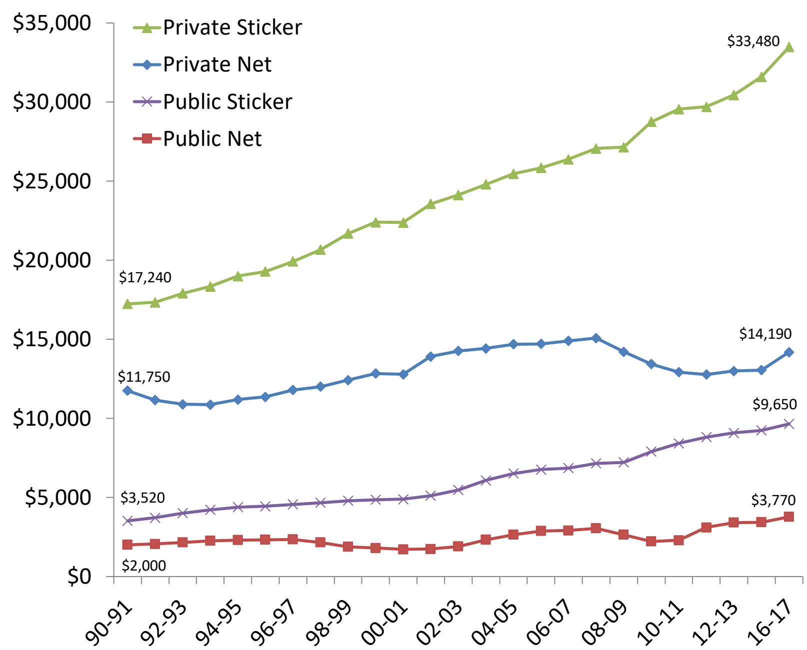
July 5, 2018

Abstract

This paper evaluates the role of rising income inequality in explaining observed growth in college tuition. We develop a competitive model of the college market in which college quality depends on instructional expenditure and the average ability of admitted students. An innovative feature of our model is that it allows for a continuous distribution of college quality. We find that observed increases in US income inequality can explain more than the entire observed rise in average net tuition since 1990 and that rising income inequality has also depressed college attendance.

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. We thank Richard Rogerson, Grey Gordon, Sebastian Findeisen, Lance Lochner, Betsy Caucutt, and Todd Schoellman for helpful comments. Corresponding author: Jonathan Heathcote. Mailing address: Federal Reserve Bank of Minneapolis, Research Department, 90 Hennepin Ave., Minneapolis MN 55401, USA. Tel: 612 204 6385. Email: heathcote@minneapolisfed.org

Figure 1: College Tuition and Fees ($2016)



Source: College Scorecard, IPEDS. This graph plots sticker and net tuition (in-state) and fees for public four-year and private nonprofit four-year colleges.

1 Introduction

The average cost of college tuition in the United States has been rising much faster than general inflation for decades (Figure 1), and paying for college has become a major concern for households with children. Policymakers worry that rising tuition costs may put a college education out of reach for high-ability children from low-income households. Given these concerns, it is important to understand what is driving up tuition. In this paper, we evaluate the hypothesis that rising income inequality has been a key driver of rising tuition.

This hypothesis is motivated by the fact that colleges in the United States draw their students disproportionately from relatively high-income households (see, for example, Figure 4 in Chetty et al., 2014). Rapid income growth at the top of the income distribution in recent decades has increased these households' willingness to pay for high-quality colleges. Lower-income households have experienced much weaker income growth over the same period, but this has likely not had a fully offsetting negative impact on college demand, given that few children from such households have ever attended college.

Predicting the impact of increasing college demand on college pricing requires modeling the college market. The model we develop follows the existing literature in recognizing two determining factors in the quality of a college education. The first is the amount of instructional resources devoted to each student. The second is the average ability of the student body, which could be interpreted as capturing average IQ or college preparedness. Schools with higher average student ability might be more attractive to college applicants for two reasons: (i) they offer better prospects for learning from peer students, and (ii) they offer social and professional connections to people who are likely to be successful postgraduation. To the extent that student ability is an important and a relatively inelastic input in producing college quality, increased demand for college will drive up equilibrium (quality-adjusted) tuition and not simply lead to an increase in the supply of high-quality college spots.

Households in our model differ with respect to household income and the ability of the household child. Colleges can observe both income and ability (e.g., by observing test scores) and in principle can price discriminate in both dimensions. Households face tuition schedules for colleges of different quality levels and decide whether to send their child to college and, if so, to which quality of college.

On the supply side, the technology for producing college quality is a constant returns to scale function of instructional expenditure per student and average student ability. There is also a fixed cost for creating each college slot. An important feature of our model, and one that is new relative to the existing literature, is that we allow for a continuous distribution of college quality.

We assume that colleges seek to provide any given value of education at the lowest possible cost or, equivalently, that they profit maximize. Colleges have no market power and thus take equilibrium tuition schedules as given. Each college chooses a quality level at which to enter, and conditional on a chosen quality level, seeks to deliver that quality as cheaply as possible by optimally balancing resource spending versus the ability composition of the student body.

As in other "club good" models, the characterization of a competitive equilibrium is complicated by the fact that club members (students) are both consumers and inputs into production, which implies a large number of market-clearing conditions. In particular, for each college quality level, the number of students demanding college spots and the ability composition of those students must be consistent with colleges' choices about the number and composition of students to "employ" as quality-producing inputs.

Given this complication, all the existing literature in the club good tradition assumes a very small number of different college quality types. The primary theoretical contribution of our paper is to allow for a continuous distribution of quality. The main practical advantage of doing so is that we can compare the equilibrium model distribution of college characteristics to US data, which include thousands of different colleges. Relatedly, our continuous distribution of college quality can change smoothly when we change income inequality or other drivers of college demand. Modeling a continuous quality distribution also has theoretical advantages, which we discuss below when contrasting our model to other important papers in the literature.

While the model features a continuous distribution of college quality and thus a continuum of market-clearing conditions and prices, it is nonetheless quite tractable. Colleges offer lower tuition to high-ability students, internalizing that such students contribute more to college quality. We prove that this ability discount is linear in ability. There is no equilibrium price discrimination by income: any such discrimination would present an opportunity to profitably skim off high-income households. Equilibrium tuition increases with college quality, which implies a natural pattern of sorting: holding ability fixed, higher-income students match in a positive assortative fashion with higher-quality schools. Combining these insights, we show that it is possible to solve for equilibrium by iterating across the quality distribution: at each quality level (i) the density of college spots satisfies total demand, (ii) baseline tuition is such that colleges make zero profits, and (iii) the tuition discount per unit of ability equates the average ability of students wishing to attend with the average ability of students that colleges want to admit.

In the first part of the paper, we characterize equilibrium in closed form in a version of the model with no resource inputs in producing college education, two ability types, and a uniform distribution for household income. We use this closed-form example to gain intuition about what determines equilibrium college prices and the distribution by quality of college spots in a club good environment and to gain insight about how these objects vary with income inequality. The comparative statics are striking. In particular, changing income inequality has absolutely no impact on the equilibrium allocation of households across colleges of different qualities and only changes equilibrium tuition pricing.

This result motivates the second part of our paper, in which we calibrate a richer version of the model and use it to explore the role of rising income inequality and other factors in driving observed changes over time in college tuition, college attendance, and the distribution of college quality. In the richer model, we explicitly model a distinction between public and private colleges. This distinction is important from a quantitative standpoint because historically, public schools have charged much lower tuition. We assume that colleges receive subsidies per student admitted and that the subsidy value is larger for public schools. The quid pro quo is that public schools face two additional constraints that do not apply to private schools. One is that they must keep quality above a certain threshold, and the second is that they face a cap on average tuition. Given this model, we show that public schools dominate the market in the middle of the quality distribution.

We set preference parameters so that the model replicates both out-of-pocket spending on college education and graduation rates. The quality and tuition thresholds for public schools are set so that — given the estimated subsidy advantage enjoyed by the public sector — we replicate both the market share of public schools and the average net tuition differential between public and private schools. We set the dispersion in student ability to replicate a measure of dispersion in within-college tuition. A key input in our calibration of the benchmark model is the joint distribution of household income and ability. We estimate a Pareto lognormal distribution for the income distribution, using data from the Survey of Consumer Finances, and use evidence from the National Longitudinal Survey of Youth (NLSY) to discipline the correlation between income and ability (proxied by Armed Forces Qualifying Test [AFQT] scores). The calibrated model generates a distribution of college tuition across colleges that is similar to that observed empirically and also replicates observed positive correlations between tuition on the one hand and measures of student ability and family income on the other.

Our key quantitative experiment is to explore the implications of the change in the US household income distribution between 1990 and 2016 for the pattern of college attendance, the distribution of college quality, and the shape of equilibrium tuition schedules. Over this period, we find evidence of both a general increase in income dispersion and a significant fattening of the right tail of the distribution. These changes can account for several key features of the data. First, rising income inequality drives up average net tuition, especially at private schools. Second, rising income inequality can account for a widening gap between average sticker tuition and average net tuition paid, reflecting increasingly large institutional discounts for desirable (high-ability) applicants (see Figure 1).

Finally, we conduct a decomposition exercise to quantify the relative roles of various potential drivers of observed changes in college attendance and college tuition. This exercise indicates that changes in the preference for college and changes in income inequality have been the most important factors driving up average net tuition, while rising subsidies have been an important countervailing force. Rising graduation rates are primarily driven by growth in average household income and by rising subsidies, while widening income inequality has depressed college attendance.

Related Literature: There is a large existing literature that models the college market. It has long been recognized that an important distinctive feature of this market is that students are both consumers of a college education and inputs into its production (see, for example, Rothschild and White 1995). Several papers in the literature build on the influential contribution of Epple and Romano (1998). In this class of models, there is a small number of colleges, which can be justified by positing large economies of scale. However, equilibrium existence problems typically arise when there are only a few college clubs, each of which is large relative to the size of the economy. These existence problems are discussed by Ellickson et al. (1999) and Scotchmer (1997) and have to do with the fact that when clubs are optimally large relative to the economy, partitioning the population into an integer number of optimally sized groups is typically not possible. Because of these problems, Epple and Romano (1998) are forced to focus on approximate equilibria. They note that one could solve the existence problem by allowing for "constant costs of schooling," which would "lead to an infinite number of schools serving infinitely refined peer groups." They note that while such a model "is extremely interesting, it is quite complex and not yet tractable" (p. 59). This infinitely-refined peer group model is the one we solve.

A second problem with assuming only a small number of competing colleges is that in such an environment, the natural model for competition is strategic oligopoly. Each college would then choose a pricing (or quality) strategy, where each strategy specifies best responses given the strategies of its competitors. Instead, most of the existing papers in the literature assume that each college takes competitors' prices (or students' willingness to pay) as given when choosing its own price. We are able to side step the difficult task of modeling strategic interactions between colleges: price-taking is the natural assumption in our competitive setting in which colleges are all small.

Caucutt (2001, 2002) takes a different approach. In her model there is a small number of different college types but a large number of colleges of each type so that each college club is small relative to the economy as a whole and there are no equilibrium existence problems. Households buy probabilities of attending different colleges, building on Cole and Prescott (1997). One limitation of her approach is that she assumes only two different income levels, which implies a very small number of different school types in equilibrium. In contrast, we assume a continuous income distribution, and we do not need to introduce lotteries (which we do not observe in practice) in order to ensure type-independent equilibrium allocations.

Two important recent papers that model the college market are Epple et al. (2017) and Fu (2014). Both papers structurally estimate rich models. Neither paper is focused on exploring the drivers of rising college tuition. In addition, there are important differences between these papers and ours in terms of how college supply is modeled. In particular, distinctive features of our model are that we allow for entry in the college market, we impose no constraints on pricing, and we assume cost minimization.[1]

There is a set of papers that explores the potential drivers of rising college tuition. Gordon and Hedlund (2017) consider various possible factors within a variant of the model in Epple et al. (2017). They find that rising financial aid is the most important factor that is pushing up tuition. This finding is in stark contrast to our model, in which more generous college subsidies lower tuition. In their model, a single monopolistic college seeks to maximize the quality of education per student enrolled. When more public aid increases students' ability to pay, this monopolist responds by increasing spending on quality-increasing inputs. In our model, more public aid induces more marginal low-income households to seek to enter college. The college market responds by expanding at the low-quality end of the distribution, thereby driving down average tuition. Jones and Yang (2016) argue that rising college tuition reflects service sector disease: productivity in higher education is assumed to be constant, but the cost of college professors continues to rise, reflecting productivity growth and a rising college wage premium in the rest of the economy. Thus, in their model rising income inequality plays a supply-side role in driving up the cost of college. We explore the role of rising supply-side costs and find them to play a relatively small role in explaining tuition trends relative to the role of changes on the demand side.

[1] Epple et al. (2006) assume that colleges maximize quality, and impose exogenous caps on the maximum tuition they can charge. Fu (2014) assumes colleges maximize a weighted sum of average student ability and a quadratic function of net tuition. In her model, colleges cannot price discriminate by ability.

There is a strong positive empirical correlation between family income and college attendance and, conditional on attendance, a strong correlation between family income and proxies for college quality (see, for example, Belley and Lochner 2007; Chetty et al. 2017; and a recent column in the *New York Times*).[2] Our model predicts similar correlations. An important finding of Belley and Lochner (2007) is an increase over time in the effect of family income on college attendance (controlling for ability). In particular, the biggest increases in enrollment between the 1979 and 1997 waves of the NLSY are for households in the top two income quartiles but the bottom two AFQT quartiles (see their Figures 2a and 2b). Chetty et al. (2017, Table 2) report changes in college attendance patterns from 2000 to 2011. For nonprofit four-year schools, the share of enrollment from households in the bottom 60 percent of the household income distribution declined over this period at selective and highly selective private and public universities. Comparing steady states calibrated to replicate observed growth in net tuition and graduation rates between 1990 and 2016, we find that the model delivers changes in enrollment patterns similar to those reported by Belley and Lochner (2007) and Chetty et al. (2017). In particular, income becomes a more important predictor of college attendance, relative to ability, with large increases in model college enrollment from low-ability households in the top half of the income distribution.

2 Model

Households: The economy is populated by a continuum of measure 1 of households, each containing a parent and a college-age child. The baseline model is static, and within each household, the parent and child operate as a single decision-making unit.

Households are heterogeneous with respect to income y and a characteristic of the child that will determine their potential contribution to college quality. We label this characteristic ability a_i and assume ability is drawn from a discrete distribution indexed by $i \in [1, ..., I]$ where a_1 denotes the lowest and a_I the highest ability levels. Let μ_i denote the corresponding population shares, with $\sum_{i=1}^{I} \mu_i = 1$. Conditional on ability, household income is continuously distributed, where the cumulative distribution $F_i(y)$ is potentially ability-type specific.

[2] Gregor Aisch, Larry Buchanan, Amanda Cox, and Kevin Quealy, "Some Colleges Have More Students from the Top 1 Percent than the Bottom 60. Find Yours," TheUpshot, *New York Times*, January 18, 2017, https://www.nytimes.com/interactive/2017/01/18/upshot/some-colleges-have-more-students-from-the-top-1-percent-than-the-bottom-60.html.

Each household chooses whether to send their child to college, and if so, to which type of college. Colleges differ by quality q. Households enjoy utility from a nondurable consumption good c and from college quality. The utility function is

$$u(c, q) = \log c + \varphi \log(\kappa + q), \tag{1}$$

where the preference parameters φ and κ are common across households. If $\kappa > 0$, then utility is bounded below for households not attending college, for whom $q = 0$.

Interpreting this utility function at face value, households are willing to pay for college because it delivers a direct consumption value. Alternatively, the same utility specification can be motivated as reflecting a setting in which college is valued indirectly as an investment technology that can increase child earnings. We will discuss this alternative interpretation further in Section 5.3.

Households take as given tuition schedules $t(q, a_i)$ that specify the out-of-pocket tuition charged by colleges offering quality q. The dependence of tuition on ability reflects the fact that different ability types are differentially attractive to colleges, and thus colleges will price discriminate.[3] The household problem, for a household with income y and ability type i, is

$$\begin{gathered} \max_{c \geq 0, q \geq 0} u(c, q) \\ s.t. \\ c + t(q, a_i) = y - \mathbb{1}_{\{q>0\}}\omega. \end{gathered}$$

If the household sends the child to college, so that $q > 0$, then it pays a fixed cost ω, corresponding to the household earnings that are forgone when the child is in college instead of at work. Let $c^i(y)$ and $q^i(y)$ describe the decision rules that solve this problem.

Colleges: Colleges can enter and supply college spots at any quality level. The technology for producing college spots is constant returns to scale, where quality depends on the average ability of the student body and expenditure (per student) e on quality-enhancing goods and services. Each college admits a continuous mass of students. Let η^i denote the

[3] In equilibrium, there is no price discrimination by income, a result we discuss later.

fraction of these students that are of ability type i. Quality (per student) is

$$q = \left(\sum_{i=1}^{I} \eta^i a_i \right)^{\theta} e^{1-\theta}, \tag{2}$$

where θ is a parameter that determines the relative importance of average student ability versus goods inputs in producing quality. Note that this production technology embeds an assumption that any mix of enrolled students with the same average ability makes an identical contribution to college quality.

Colleges must also pay a fixed resource cost ϕ per student enrolled, which captures administration and other costs that do not directly enhance quality. The fact that the technology for supplying college spots is constant returns to scale supports the existence of an equilibrium in which all colleges are small, in the sense that they enjoy no pricing power. The logic is that each college competes against other colleges offering identical quality.[4] Colleges observe the ability and income of applicants. They seek to maximize profits or, equivalently, to provide a given market value of education at the lowest possible cost.[5]

Suppose a college has decided to supply college education at quality level q. The input mix subproblem for supplying mass one spots at quality q is

$$\begin{aligned}
&\max_{\eta^i \geq 0, e \geq 0} \left\{ \sum_i \eta^i t(q, a_i) - e - \phi \right\} \\
&s.t. \\
&q = \left(\sum_i \eta^i a_i \right)^{\theta} e^{1-\theta} \\
&\sum_i \eta^i = 1.
\end{aligned}$$

Let $\{\eta^i(q)\}$ and $e(q)$ denote the values for $\{\eta^i\}$ and e that solve this problem, and let $\pi(q)$ denote corresponding profit per student. Given tuition schedules $t(q, a_i)$, colleges will optimally supply zero mass of college spots at qualities q where $\pi(q)$ is negative, will be indifferent about the mass of spots to supply if $\pi(q) = 0$, and will want to supply an infinite mass of spots if $\pi(q)$ is strictly positive.

[4] Even if a college had a monopoly at a given quality level, it would still face near-identical competitors given a continuous quality distribution and thus enjoy no pricing power.

[5] Some other papers in the literature assume that colleges seek to maximize college quality. In a competitive environment in which colleges take tuition schedules as given, quality maximization would imply a degenerate college quality distribution, with all colleges operating at the highest feasible quality level.

Let $\chi(Q)$ denote the measure of college places in colleges of quality $q \in Q \subset \mathbb{R}^+$. This is a key equilibrium object. In contrast, the size distribution of colleges within any given quality level is indeterminate, given our constant returns to scale quality production function.

Definition of Equilibrium: An equilibrium in this model is a measure χ and functions $t(q, a_i)$, $c^i(y)$, $q^i(y)$, $\eta^i(q)$, $e(q)$, and $\pi(q)$ that satisfy the following conditions:

1. Given $t(q, a_i)$, the household choices $q^i(y)$ and $c^i(y)$ solve the household's problem for all y and $i \in I$.

2. Given $t(q, a_i)$, the college input choices $\eta^i(q)$ and $e(q)$ solve the college's problem for all $q > 0$, and $\pi(q)$ is the associated profit per student.

3. Zero profit condition: For all $Q \subset \mathbb{R}^+$, $\int_Q \pi(q) d\chi(q) = 0$ and $\pi(q) \leq 0$ for all $q \in Q$.

4. The goods market clears:
$$\sum_{i=1}^{I} \mu_i \int_0^\infty c^i(y) dF_i(y) + \int_0^\infty e(q) d\chi(q) + (1 - \chi(0))(\omega + \phi) = \sum_{i=1}^{I} \mu_i \int_0^\infty y dF_i(y). \tag{3}$$

5. The college markets clear. For all i and $Q \subset \mathbb{R}^+$,
$$\mu_i \int \mathbb{1}_{\{q^i(y) \in Q\}} dF_i(y) = \int_Q \eta^i(q) d\chi(q), \tag{4}$$
where $\mathbb{1}_{\{.\}}$ is an indicator function.

Condition (3) here is the zero profit condition that follows from free entry and perfect competition. It states that profits are not strictly positive at any quality level and that average profits are identically zero over any quality values at which a positive measure of college spots are supplied. Condition (4) is the goods market-clearing condition. In addition to the variable cost e, each student attending college also consumes a fixed resource cost $\omega + \phi$. The college market-clearing conditions are described in condition (5): for each ability type and for each possible set of college qualities, the number of students that wish to attend a college in that quality set must equal the corresponding number of spots supplied. The fact that there are many such market-clearing conditions reflects the club good setting.[6]

[6] Ellickson et al. (1999) prove the existence of equilibrium in a similar setting.

2.1 Equilibrium Characterization

With multiple ability types, characterizing equilibrium in principle requires solving for a separate tuition schedule for each ability type, with the property that the college market clears for each quality-ability combination. However, it can be shown that the college tuition schedule is linear in ability, which greatly simplifies equilibrium characterization.

Proposition 1

Any equilibrium in this model can be supported by a tuition schedule that is linear in ability, that is, a tuition schedule that takes the form

$$t(q, a_i) = b(q) - d(q)(a_i - a_1).$$

The quality-dependent constant $b(q)$ defines the baseline tuition charged to the lowest ability type, while $d(q)$ denotes the tuition discount per unit of ability. The logic underlying Proposition 1 is that college quality depends only on the average of the student body (in addition to expenditure e), and thus only average student ability is priced in equilibrium. The tricky part of this proof is when only a subset of ability types attend colleges of a given quality, there may be a set of tuition schedules consistent with the same equilibrium. We construct a linear tuition schedule within this set.

A linear tuition function ensures that the tuition revenue from admitting any set of students is linear in the average ability of the students in that set. Thus, we can rewrite the college problem as

$$\begin{aligned}
&\max_{\bar{a},e} \{b(q) - d(q)(\bar{a} - a_1) - e - \phi\} \\
&s.t. \\
&\ q = \bar{a}^{\theta} e^{1-\theta} \\
&a_1 \leq \bar{a} \leq a_I,
\end{aligned}$$

where $\bar{a} = \sum_i \eta^i a_i$ denotes the average ability of students admitted.[7] The implied first-order condition is

$$\frac{d(q)}{1} = \frac{\theta \frac{q}{\bar{a}}}{(1-\theta)\frac{q}{e}}. \tag{5}$$

The left-hand side is the ratio of the price of a marginal increase in average ability relative

[7] Note that $\bar{a}$ is a continuous choice variable, even though a_i takes a discrete set of values because each college enrolls a continuum of students and can vary average student ability smoothly by varying the enrollment shares η^i.

to the price of a marginal increase in e. The right-hand side is the corresponding ratio of marginal products. If the equilibrium tuition schedule declines steeply with ability ($d(q)$ is large), then colleges will choose a high ratio of instructional inputs relative to average student ability.

With the college problem reformulated this way, we can simplify the college market-clearing condition, replacing eq. 4 with the following two conditions:

$$\begin{aligned}\chi(Q) &= \sum_i \mu_i \int \mathbb{1}_{\{q^i(y)\in Q\}} dF_i(y) \qquad \forall Q \subset \mathbb{R}^+, \\ \int_Q \bar{a}(q) d\chi(q) &= \sum_i \mu_i \int \mathbb{1}_{\{q^i(y)\in Q\}} a_i \, dF_i(y) \qquad \forall Q \subset \mathbb{R}^+.\end{aligned}$$

The first condition states that the measure of students in any quality set Q is consistent with student attendance choices. The second equates the average student ability demanded by colleges producing in quality set Q to the average ability of the students choosing to supply to quality set Q.

Other properties of equilibrium tuition schedules are more immediate. First, for any two quality levels in positive supply, it must be the case that strictly higher quality translates into strictly higher tuition, conditional on ability. The reason is that if the higher-quality school were no more expensive, then no students would choose the lower-quality college. Second, tuition must be declining in student ability (i.e., $d(q) > 0$). Otherwise, colleges would be able to increase tuition and profit by admitting a higher average ability student body and simultaneously reducing instructional expenditure e to leave quality unchanged. Third, equilibrium tuition must be independent of household income, holding fixed student ability. Suppose, to the contrary, that all colleges were charging high-income students more than equally able low-income students. Any single college would then be able to increase profit by skimming off the high-income students.

We now turn to households' choices about college attendance and quality. Conditional on ability, the optimal choice for quality is increasing in household income y. This is because college quality is a normal good and because equilibrium tuition, as just discussed, is independent of income. This property simplifies equilibrium computation because it means that moving up the college quality distribution, college spots for each ability type will be filled in a strictly ordered fashion by income, with the highest-quality colleges taking students from the top of the income distribution. If colleges must cover positive fixed costs (ϕ is positive), then average tuition (weighted by ability shares) must be positive at any quality level, and thus equilibrium tuition $b(q)$ for the lowest ability students must be positive. Given a positive reservation utility parameter κ, the marginal utility from college quality is bounded below, and thus households with lowest-ability children and sufficiently low income will prefer not to attend college.

Our competitive model implies an absence of income-based price discrimination. Is this prediction counterfactual? In practice, a portion of institutional financial aid is labeled "need-based," which might be interpreted as income-based price discrimination. However, many colleges promise to meet the "full demonstrated financial need" of admitted students but are "need-sensitive" at the admission stage. Among applicants who will need significant aid, these schools presumably admit only the strongest. Thus, aid that they describe as "need-based" actually has a "merit-based" component. Similarly, in the equilibrium of our model, a typical model college will contain two types of student: low-ability students with relatively high income, and high ability students with relatively low income. Because low ability students pay higher tuition in equilibrium, this selection pattern leads to the appearance that low-income students enjoy tuition discounts.

Fillmore (2016) argues that colleges exploit information on FAFSA forms to price discriminate by income. He shows that higher family income translates into smaller offered tuition discounts, even after controlling for observable proxies for ability (ACT scores and high school GPAs). This pattern emerges primarily at the most selective schools, suggesting that such colleges have some pricing power. One way to reconcile income-dependent pricing with a competitive framework would be to assume that student ability is not perfectly observable and that family income is a useful signal of ability. For example, if two college applicants have the same SAT score, but one comes from a much poorer family, then one might infer that the lower-income student actually has greater true ability. A second way to introduce income-dependent pricing in a competitive framework is to posit that students benefit from having peers from diverse backgrounds, which would make low-income students attractive to colleges that draw predominantly from the top of the income distribution.

2.2 Pareto Efficiency

Given that there is an externality in the form of a peer group effect in this model, one might conjecture that the competitive equilibrium is not efficient and thus that there is a rationale for government intervention. Contrary to this intuition, the competitive equilibrium is in fact Pareto efficient. A crucial assumption is that clubs are competitive price takers. This implies that the peer group effect is a "local" externality within a college and is correctly priced in the competitive market: higher-ability students are charged lower tuition. As argued in Ellickson et al. (1999), a club goods economy is conceptually no different from a non-club-goods economy in the sense that (type-specific) club memberships can be treated as ordinary goods traded in a competitive market. We summarize the discussion in the following proposition.

Proposition 2
A competitive equilibrium in the model described is Pareto efficient.

The proof here closely follows the standard proof of the First Welfare Theorem. The result hinges critically on the assumption that households care about quality directly and only derive utility from their own child's college quality, so college is a pure consumption good. High income households will spend a lot on tuition at expensive schools – even if their children are low ability – because their marginal utility from non-college consumption is low. Thus, they have a high willingness to pay for the experience and prestige associated with high quality colleges, or for gaining access to a more attractive pool of potential spouses and professional connections. Similarly, if poor households choose not to send their children to college that will simply reflect the fact that those households prefer to spend their limited income on consumption goods.

In Section 5.3 we discuss an alternative interpretation of the model in which the value of college has an investment component, in which case allocations are not efficient.

3 A Closed-Form Example

Before calibrating the model described above, we first show that in one special case, equilibrium allocations can be characterized in closed form. This example is useful because it clearly illustrates that the club good nature of the college market has important implications for college pricing and for the effects of changes in income inequality on the allocation of students to colleges and the tuition they pay.

The special case is one in which $\theta = 1$, so average student ability is the only determinant of college quality. In addition, $\omega = \phi = 0$, so there are no fixed resource costs associated with attending college. The preference parameter $\varphi = 1$. There are two ability types, which we denote a_l and a_h, and half the population is of each type. The income distribution is independent of ability and uniform: $y \sim \mathcal{U}(\mu_y - \frac{\Delta_y}{2}, \mu_y + \frac{\Delta_y}{2})$ where μ_y denotes average income and Δ_y defines income dispersion. Let $\Delta_a = a_h - a_l$ denote ability dispersion and $\mu_a = \frac{a_h + a_l}{2}$ denote average ability.

Note that, given $\theta = 1$, the production function implies that $q = \bar{a}$. Note also that the support of possible college qualities is $[a_l, a_h]$.

Proposition 3

Under the parameterization described above, the model has a unique competitive equilibrium in which the measure of college spots by quality and tuition schedules are described, respectively, by

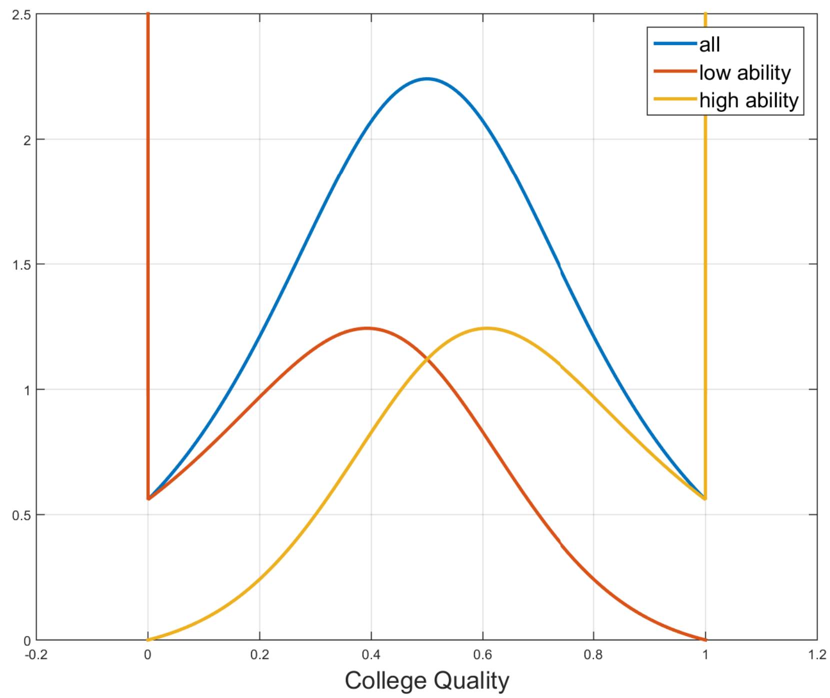
$$\begin{aligned} \chi(Q) &= \frac{2}{\Delta_a} \frac{2}{(4+\pi)} \int_Q \left[\left(\frac{a_h - q}{\Delta_a} \right)^2 + \left(\frac{q - a_l}{\Delta_a} \right)^2 \right]^{-2} dq \;\; \forall Q \subset (a_l, a_h) \\ \chi(a_h) &= \chi(a_l) = \frac{2}{4+\pi} \end{aligned}$$

$$t(q, a_i) = \mu_y \frac{q - a_i}{\kappa + q} \left[1 - \frac{2}{4+\pi} \frac{\Delta_y}{\mu_y} \arctan\left(2 \frac{(\mu_a - q)}{\Delta_a} \right) \right]$$

where π is the mathematical constant and arctan *is the inverse tangent function.*

The equilibrium allocations have several interesting properties. First, the quality distribution of college spots is independent of the income distribution parameters μ_y and Δ_y and is also independent of the preference parameter κ. Given perfect assortative matching from income to quality, this result implies that the equilibrium college quality choice for a household depends only on the household *rank* in the income distribution and their child's ability. The quality distribution of college spots has two mass points at the lowest and highest possible quality colleges, $q = a_l$ and $q = a_h$. In between these values, the distribution is continuous, symmetric, and single-peaked. The lowest-quality schools are filled with low-ability students drawn from the bottom of the income distribution, while the highest-quality schools are

Figure 2: College Distribution

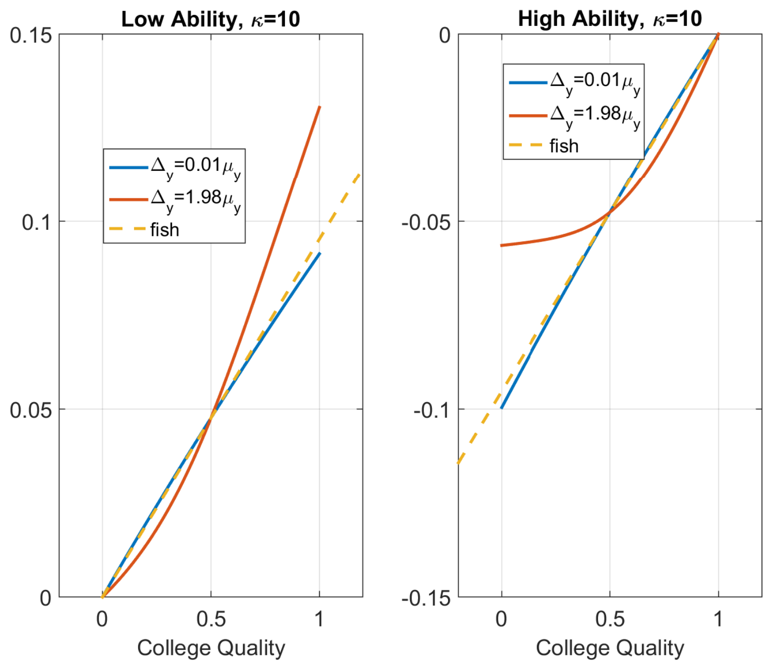


filled with high-ability students drawn from the top. Outside these income ranges, students attend mixed-ability schools.

In contrast, the income distribution parameters μ_y and Δ_y do appear in the equilibrium tuition functions. Tuition is an increasing but a nonlinear function of quality. It is easy to check that $t(q, a_l) \geq 0$ and $t(q, a_h) \leq 0$ for all $q \in [a_l, a_h]$. The tuition ability discount $d(q) = \frac{t(q,a_l)-t(q,a_h)}{\Delta_a}$ is increasing (decreasing) in income dispersion Δ_y for $q \geq \mu_a$ ($q \leq \mu_a$). Figures 2 and 3 plot $\chi(q)$ and $t(q, a_i)$ for an example in which $a_l = 0$, $a_h = 1$, $\mu_y = 1$, and $\kappa = 10$.

The key feature of our club good model is that households can choose only one quality of college: if they choose to "sell" their child's ability to Harvard, they must also choose to "buy" the average ability of Harvard students. We now briefly contrast the club good model with a (counterfactual) alternative in which student ability is modeled as a conventional good that can bought and sold on a centralized market. In this alternative model, a poor but high-ability household can sell their child's ability to a high average ability school and

Figure 3: College Tuition Schedule



buy a cheap education from a low average ability school. Ability has a fixed per unit price p, and tuition can be defined as the difference between ability bought minus ability sold, $t_{LP}(q, a_i) = pq - pa_i$, where LP denotes linear pricing.

The differences between the baseline club good model and the linear pricing model are stark. In the linear pricing model, $p = \frac{\mu_y}{\mu_a+\kappa}$, and thus a mean-preserving increase in income inequality (an increase in Δ_y) has no impact on tuition prices but instead increases dispersion in the distribution of quality demand. Both properties reflect the fact that quality demand is linear in income. In the club good model, in contrast, the comparative statics are exactly the opposite. Now an increase in Δ_y has *no* impact on the distribution of quality demand but *does* change equilibrium tuition prices.

Why is the distribution for college quality χ completely insensitive to income inequality in the club good model? A partial intuition is that the support of the college quality distribution cannot expand in our environment: the bounds are always $q = a_l$ and $q = a_h$. Clearly, in order for households at the extremes of the income distribution to choose feasible values for quality, tuition schedules must move when income inequality is increased. The fact that the richest low-ability households are now richer increases the relative demand from low-ability households for high quality colleges. In equilibrium, a rise in low-ability tuition at high-quality colleges induces these richer households to leave their quality choices unchanged. Similarly, rising income inequality leaves the poorest high-ability students poorer. A rise in high-ability tuition at low-quality colleges induces these households not to downgrade college quality.

4 Quantitative Application

Public versus Private Colleges: We now turn to our quantitative application. Here we extend the model in order to differentiate in the theory between public and private universities. American public universities receive larger direct government support than their private competitors. The *quid pro quo* is that public universities are expected to keep tuition low.

Formally, we assume that in addition to choosing quality q, each college also chooses a business model $j \in \{1, 2\}$, where $j = 1$ denotes public status and $j = 2$ private. Colleges then receive a subsidy s_j per student admitted. Our calibration will feature $s_1 > s_2$. This larger subsidy gives public schools a competitive edge, but we assume that public schools also face two additional constraints. First, public schools face a constraint on the average tuition they charge, of the form $t(q, \bar{a}(q)) \leq \bar{T}$. This constraint effectively caps public college spending on instructional inputs and gives private colleges a competitive advantage at high-quality levels. Second, public schools face a constraint on quality, $q \geq \underline{Q}$, which gives private colleges an advantage at low quality levels. Thus, the public sector will dominate in the middle of the quality distribution. In other respects, public and private colleges are identical. In particular, all colleges profit maximize, and all operate the same production technology.

Note that, from a college's standpoint, subsidies are isomorphic to (negative) fixed costs, and thus a college only cares about the net fixed cost $\phi - s_j$. Also, in our competitive environment, it does not matter whether subsidies are paid directly to colleges or as enrollment-contingent transfers to students. In our calibration, we will measure subsidies by aggregating across all sources of non-tuition revenue, including the value of government grants to students, direct government support to colleges, and income from endowments and other sources.

4.1 Calibration

Our baseline calibration is for 2016. We will also later be interested in a calibration for 1990, which we will use to explore the drivers of changes over time in college attendance and tuition.[8] We assume that the distribution for household income (conditional on child ability) is Pareto lognormal, a parametric functional form that closely approximates the actual distribution of income in the United States (see Heathcote and Tsujiyama 2017). Thus, log household income is given by $\ln y = x_1 + x_2$, where x_1 and x_2 are independent random variables, x_1 is normally distributed with mean μ and variance σ^2, and x_2 is exponentially distributed with exponential (Pareto index) parameter α. This distribution transitions smoothly from an approximately lognormal distribution over most of the income distribution toward a Pareto distribution in the right tail. We estimate the parameters μ, σ^2, and α using microdata on log total household income from the 2016 Survey of Consumer Finances. One important strength of this survey is that households at the top of the income distribution are not underrepresented, which is important for being able to estimate the Pareto parameter α. Because we are interested in income for households who are making decisions about college, we restrict our sample to households between ages 40 and 59. The maximum likelihood estimates for σ^2 and α are 0.55 and 1.67, implying a variance of log income equal to $0.55 + 1.67^{-2} = 0.91$.

As in the closed-form example, our baseline calibration assumes two ability types. In Appendix 7.2.2, we show that introducing more ability types has a negligible impact on the distribution of enrollment and tuition.[9] We assume that the income distributions conditional on ability are both Pareto lognormal, with the same estimated values for σ^2 and α. To allow for correlation between household income and child ability, we index the level parameter μ_i by ability $i \in \{h, l\}$. To estimate how household income varies by child ability, we turn to the 1997 NLSY and use AFQT scores as a proxy for child ability. We rank households by these scores and set the ratio μ_h to μ_l to replicate the ratio of average family income for households with children in the top versus the bottom half of the AFQT score distribution, which is 1.49.

College attendance in the model is sensitive to the reservation utility parameter κ. In

[8] See the Data Appendix for more details on the construction of the statistics used in calibration and model-data comparisons.

[9] We develop a specialized computational algorithm for the two-ability-type case, which is much faster than our general algorithm, which can be used for any number of ability types.

Table 1: College-Level Statistics (Public vs. Private)

		Public	Private
(1)	Share of students	0.696	0.304
(2)	Sticker tuition	9,650	33,480
(3)	Federal and state grant aid	3,199	2,893
(4)	Institutional aid	2,681	16,397
(5)	Net tuition	3,770	14,190
(6)	Forgone earnings	20,040	20,040
(7)	Room and board	10,440	11,880
(8)	Instructional spending	10,221	17,567
(9)	Student services	1,660	4,553

Notes: The full cost of attending college for a low-ability student is (2) + (6) − (3).
The average cost of attending college for all students is (2) + (6) − (3) − (4) = (5) + (6).
The total per student subsidy net of fixed costs is $s_j - \phi = E[e|_j] - E[t^i(q)|_j]$ =(8) + (9) − (5).

2016, 37 percent of individuals aged 25-34 reported having at least a bachelor's degree, and this is the graduation rate we target.[10]

The remaining model parameters are set to replicate various moments involving college costs, tuition, and financial aid for the universe of nonprofit four-year private and public colleges.[11] Data on tuition, aid, and costs are from the College Board and the National Center for Education Statistics (NCES). Table 1 describes the empirical statistics used as inputs for this part of the calibration.

The preference parameter φ is set to replicate average college tuition. We focus on net tuition (row 5 of Table 1), defined as sticker tuition minus federal, state, and institutional grant aid. This reflects the average amount students pay out of pocket and is therefore a good gauge of the strength of the preference for college. Weighted by public versus private graduate shares, average net tuition for the 2016-2017 academic year was $\$6,942$. Forgone earnings ω is an important additional component of the cost of attending college. We set ω equal to 40 weeks of median weekly earnings for full-time workers aged 16 to 24, which was $\$20,040$ in 2016 (Current Population Survey). We exclude room and board from our measure of college prices, on the grounds that similar costs apply irrespective of college

[10]Our model does not differentiate between college enrollment and college graduation. Hendricks and Luekhina (2017) argue that college entrants can fairly accurately predict whether or not they will graduate.

[11]Around one-tenth of individuals with a bachelor's degree are graduates from for-profit colleges, which could be conceptualized as low-quality private schools in the context of our model. However, because of data limitations, we exclude the for-profit sector when computing the statistics targeted in our calibration.

attendance.[12] While our model is static, in practice college graduates only spend a fraction of their lifetime in college. In 2014, there were 8.26 million students enrolled in four-year schools, out of a total US population of 318.9 million, implying an aggregate attendance rate of 2.59 percent.[13] Private consumption per capita in 2016 was $\$39,417$ (national income and product accounts). Thus, aggregate out-of-pocket tuition spending on four-year colleges was $2.59 \times \frac{6{,}942}{39{,}417} = 0.46$ percent of total consumption.[14] We set φ to replicate this ratio in the model.

We now turn to the college quality production function. We need to identify fixed costs net of subsidies, $\phi - s_j$. In equilibrium, colleges make zero profits. We can therefore estimate the net fixed cost $\phi - s_j$ for each sector j by taking the difference between average net tuition paid and average variable resource spending. We identify variable resource spending e with the sum of the NCES expenditure components "instruction" and "student services." Fixed cost expenditures ϕ are the residual categories "administration" and "maintenance of plant."Instruction and student services spending per student averaged $\$22,120$ at private four-year schools, implying negative net fixed costs $\phi - s_2$ of $-\$7,930$. The analogous estimates for expenditure and net fixed costs for public schools are $\$11,881$ and $-\$8,111$. The fact that net fixed costs $\phi - s_j$ are negative indicates that both private and public schools receive significant nontuition revenue, with public schools receiving more direct state government support and private schools relying more heavily on endowment income and private gifts.

Next, we turn to the ratio between ability levels, a_h/a_l, and the share parameter θ, which defines the relative importance of average student ability versus instructional spending in determining college quality. These parameters jointly determine the importance of the club good feature of the model and have similar effects on model observables. In particular, the extent of model within-college tuition variation reflects both the size of ability differentials in the student body, a_h/a_l, and the importance of ability as an input, θ. There is no within-college tuition variation if either $\theta = 0$ or $a_h/a_l = 1$. We have limited data on within-

[12] Increased top tail income inequality will translate into greater demand for high quality accommodation and food, and thus can potentially rationalize observed growth over time in the cost of Room and Board.

[13] Given the observed 37.0% graduation rate, and assuming the U.S. economy is in steady state, this translates into each college graduate spending 7.0 percent of their lifetime in college.

[14] The share of college tuition and fees in the Consumer Price Index is larger at 1.8%. However, the CPI college category includes two-year schools and graduate and professional schools. Also the CPI tuition measures are closer to capturing sticker price than actual price paid (e.g., Schwartz and Scafidi 2004).

college tuition dispersion, but colleges do report sticker tuition and average net tuition paid. Assuming that sticker tuition reflects the cost of enrollment for low-ability students, we can pin down the ratio a_h/a_l by targeting the empirical ratio of average sticker tuition to average net tuition paid, which was $\$13,800/\$6,942 = 2.0$ in 2016.[15] Because a_h/a_l and θ are not sharply separately identified, we then simply set $\theta = 0.5$ following much of the existing literature, implying that peer effect and expenditure are equally important in delivering quality (see, e.g., Caucutt 2002 and Epple et al. 2017).[16]

We set the constraints on quality and tuition for public colleges, $\underline{Q}$ and $\bar{T}$, in order to replicate the ratio of average net tuition for public relative to private colleges and the public sector's share of college graduates. Intuitively, reducing $\bar{T}$ reduces average public tuition, while raising $\underline{Q}$ shrinks public enrollment. The calibration is summarized in Table 2.[17]

4.2 Model Predictions

Figure 4 plots the distribution of college quality under the baseline calibration.[18] Several features of the plot are worth noting. First, three quality ranges have positive mass. The bottom- and top-quality intervals are supplied by private colleges, while the middle-quality interval is public. Second, there are no very low-quality colleges: demand for such colleges is weak given the lost-earnings cost ω of attending college in combination with a positive reservation quality κ when not attending. Third, there is a fat right tail in the quality distribution, mirroring the right tail of the income distribution. Fourth, there are gaps

[15]In the model, given our two-point ability distribution and the fact that twice as many high-ability children attend college relative to low-ability children, only one-third of model college graduates pay sticker tuition. This is consistent with College Board estimates of the fraction of college students that pay sticker price. In the ratio described in the text, we reduced sticker price tuition by the size of average federal and state grant aid, under the assumption that this aid is provided equally to high- and low- ability students.

[16]The standard deviation of test scores on standardized tests cannot be used to identify the ratio a_h/a_l since test score dispersion is a choice of the test designer rather than an informative empirical moment that can be targeted.

[17]In Table 2, the parameters μ_h and μ_l are scaled so that average household income is equal to one. The parameters a_h and a_l are scaled so that average student ability is equal to one. The parameters σ^2, α, φ, and θ are independent of the scale of average income or average ability. The parameters ω, $\phi - s_j$, and $\bar{T}$ scale proportionately with average income. In addition, these parameter values reflect lifetime cost and tuition values and should be divided by the fraction of lifetime in college to be interpretable as costs per year in school. Thus, for example, the public college tuition cap is $0.038 \times \$39,417 \div 0.07 = \$21,398$.

The parameters κ and $\underline{Q}$ scale with average income (average ability) in proportion to the share of resources (ability) in quality production, so that if average income (ability) is scaled by a factor λ, then κ and $\underline{Q}$ are scaled by $\lambda^{1-\theta}$ (λ^{θ}).

[18]The quality axis is truncated at $q/\kappa = 14$.

Table 2: Calibration

Parameters	Value	Source/Target
Income distribution		
μ_h, μ_l	$-1.008, -1.406$	Income conditional on AFQT score (1997 NLSY)
σ^2	0.55	2016 Survey of Consumer Finances
α	1.67	
Preferences		
φ	0.031	Tuition share of aggregate consumption 0.46%
κ	0.033	College attendance 37%
Technology		
θ	0.5	Peer effects and expenditure equally important (Caucutt 2002; Epple et al. 2017)
a_h, a_l	$1.363, 0.637$	Ratio of sticker to net tuition equal to 2.0
ω	0.036	40 weeks' earnings for 16- to 24-year-olds (Table 1)
Policy		
$\phi - s_j$	$-0.0145, -0.0142$	Net fixed cost per Full-time equivalent student (Table 1), equivalent to $\$8,111$ and \$7930 for public and private colleges
$\bar{T}$	0.038	Public sector's share of graduates (Table 1)
$\underline{Q}$	0.12	Public relative to private net tuition (Table 1)

Figure 4: College Distribution

calibration_distribution.pdf

within the quality distribution. The first gap is just below the lower bound constraint on quality faced by public schools. Private schools could offer quality slightly below this bound at slightly lower prices, but this would imply negative profits given that they receive smaller subsidies than public schools. The second hole in the quality distribution emerges where public schools cannot make a profit because of the tuition cap they face, while private schools cannot make a profit because they are competing with similar but more heavily subsidized private schools. Finally there is a mass point in the distribution at the lowest quality public school: 6.3 percent of children attend colleges with $q = \underline{Q}$.

Figure 5 plots equilibrium tuition schedules by ability. Tuition increases with quality in a nonlinear fashion. Tuition per unit of quality is also increasing in quality. As expected, high-ability students pay less than low-ability ones at all quality levels, as their role as quality-enhancing inputs translates into "merit-based" financial aid.

Panels (A) and (B) of Figure 6 plot the level of expenditure per student by college quality and the fraction of high-ability students enrolled. Higher-quality schools spend more per

Figure 5: Tuition Schedules and Tuition per Unit of Quality

calibration_tuition.pdf

Figure 6: College Inputs

calibration_eta.pdf

student and also generally admit a student pool with higher average ability. However, Panel (C) indicates that the tuition discount per unit of ability $d(q)$ (equivalently, the equilibrium differential between tuition for high- and low- ability students) is increasing in college quality. Thus, as quality increases, the effective price of the ability input rises, inducing colleges to increase the ratio of expenditure to average student ability (Panel (D); also see eq. 5). This in turn implies that the ratio of quality to expenditure declines with quality, which explains why tuition per unit of quality is increasing in quality.

Note that as the distribution moves toward the highest-quality public schools, a spike is evident in the share of high-ability students. Here the public colleges' tuition cap is binding, which translates to an effective cap on instructional spending, $e = \bar{T} + s_1 - \phi$. Within a narrow quality range, however, public schools can increase quality without increasing spending by admitting a larger share of high-ability students (whose tuition is discounted).

Figure 7 describes how students are allocated to colleges. The blue line describes college outcomes for high-ability students from different percentiles of the household income distribution, while the red line shows the same thing for low-ability students. High-ability students below the 50.4^{th} percentile of the ability-conditional household income distribution do not attend college. Moving up through the income distribution, we see that students attend increasingly higher-quality colleges. Note that jumps in the quality allocation rule at the points in the income distribution where households are indifferent between private and public schools; these points correspond to the gaps in the $\chi(q)$ distribution. Flat spots occur in the quality allocation rule for households choosing the worst-quality public school; the flat spots correspond to the mass point at this quality value in the $\chi(q)$ distribution. Only 24.5 percent of low-ability children attend college in the model, and of this percentage, most

Figure 7: Student Allocation across Colleges

calibration_income_quality.pdf

attend public colleges or low-quality private colleges. Low-ability students who lie above the 98.9^{th} percentile of the family income distribution attend high-quality private colleges, paying a steep tuition premium (see Figure 5).[19]

Table 3 compares some moments of the calibrated model to the data. Average graduation and tuition are from Table 1. All other moments are from the College Scorecard, which is a tool provided by the US Department of Education to facilitate comparison shopping across colleges, supplemented with parental income data from Chetty et al. (2017). The College Scorecard reports various percentiles of SAT and ACT scores for admitted students. We assume these scores are normally distributed with college-specific means and variances. We then estimate the fraction of students within a college whose score lies in the top half of the national test score distribution and identify this fraction as the share of the college student body that is high ability. The College Scorecard also reports sticker tuition and fees, and the average net price of attendance for undergraduates receiving Title IV aid, which is the full cost of attendance less federal, state, and institutional grant aid.[20] Chetty et al. (2017) report average parental family income by college, constructing these estimates from federal income tax records. Parental income is averaged over 1996-2000, a five-year period in which the child (and potential college attendee) is aged 15-19.

The model replicates sector-specific graduation and average net tuition by construction. Average sticker tuition amounts (gross of federal and state grants) in the model and data are similar. In terms of family income, the model implies a realistic value for average family income for students attending public colleges, but the model's family income for private

[19]One might argue that the model prediction that low-ability students can buy admission to high quality colleges is counterfactual and thus constitutes evidence against our competitive model of the college market. However, there is evidence that large financial donations do facilitate admission at elite private colleges in the United States (see Golden 2006). For such students, sticker price tuition vastly understates the true cost of admission.

[20]This measure includes living expenses. To construct a net tuition measure that excludes living expenses, we estimate living expenses at the college level as the difference between the average annual full cost of attendance and tuition and fees. Average net tuition in our sample as measured by the College Scorecard is $\$4,904$, which is lower than average net tuition as reported by the College Board (see Table 1).

Table 3: College-Level Moments Comparison

First Moments	Model		Data (2016)	
	Public	Private	Public	Private
Graduation	25.8	11.3	25.8	11.3
Net tuition $	3,789	14,226	3,770	14,190
Sticker tuition $	12,123	27,248	9,650	33,480
Avg. family income / Mean	1.63	2.90	1.37	1.95
Fraction of high ability	0.71	0.57	0.74	0.79

Second Moments	Model	Data (2016)
Standard Deviation/Mean		
Net tuition	0.876	0.989
Sticker tuition	0.947	0.769
Avg. family income	1.572	0.509
Fraction of high ability	0.184	0.258
Correlation		
Net tuition vs. Family income	0.774	0.603
Net tuition vs. Fraction of high ability	0.435	0.218
Family income vs. Fraction of high ability	0.795	0.585

college students is higher than in the data. The fraction of students that are high ability is similar in the model and data for the public sector, but the model delivers too small a share of high-ability students in the private sector. This discrepancy reflects the existence in the model of low-quality private colleges (see Figure 4).[21]

In terms of second moments, the model generates realistic dispersion in tuition and in the share of high-ability students across schools. Dispersion in average income across schools in the model is too large relative to the data, reflecting the monotonic relationship, conditional on ability, between income and college quality.[22] The model generates realistic comovement at the college level between average household income, average net tuition, and the estimated fraction of high-ability students.

Figure 8 plots the distribution of published tuition and fees, model against data. The data here are from College Board (2016a). The model distribution is binned consistently

[21]Around one-third of non-profit private colleges do not report standardized test scores. Schools with missing scores tend to charge low tuition, and we therefore conjecture that they admit relatively few high-ability students. Thus, our estimate of the share of high ability students in the private sector is likely biased upward. In addition, one could interpret some portion of the low-quality segment in the model as capturing the for-profit sector, which is not part of our data sample and which likely admits mostly low-ability students (many for-profit schools do not require or report standardized test scores).

[22]A mechanical way to reduce the variation in average income across colleges would be to introduce heterogeneity in the preference parameter φ for college, which would blur the positive association between household income and willingness to pay.

Figure 8: Sticker Tuition Distribution

stick_distribution_all.pdf

with the data, and for the purposes of consistent measurement includes federal and state grant aid. Overall, the distributions look broadly similar. One discrepancy is at the top of the quality distribution. In the model (as in the data), there are some very rich households who are willing to pay very high tuition for very high-quality colleges. In the model, the market responds by supplying such colleges. In the data, there is a group of elite private colleges charging around $\$65,000$, but few schools charge more.[23]

Overall, our relatively simple calibrated model successfully replicates some important features of the US college market: how many people attend, attendees' average income and ability, and observed variation across colleges in terms of tuition and the typical income and ability backgrounds of their customers.

5 Understanding Changes in College Tuition

5.1 The Effect of Rising Income Inequality

We now explore the implications of changes in income inequality for college enrollment and tuition. We reestimate the income distribution parameters σ^2 and α using household income data from the 1989 wave of the Survey of Consumer Finances and find that $\sigma^2 = 0.48$ and $\alpha = 2.4$. Thus, the implied variance of log income in 1989 is $0.48+2.4^{-2} = 0.65$, compared to 0.91 in 2016. Note that most of the increase in the variance over this period is attributable to a heavier right tail in the income distribution.[24] Figure 9 plots the estimated exponentially modified Gaussian (EMG) distribution for log household income in the two years. Note that while mean log income is very similar in both years, mean level income is 24.7 percent larger

[23] However, successful parents and successful graduates at elite private colleges traditionally make financial contributions on top of paying tuition. Thus, the plot for published tuition understates top-tail inequality in resources available for instruction.

[24] Data from the online appendix to Piketty and Saez (2003) indicate a similar increase in the Pareto parameter. For example, one can estimate the Pareto parameter as average income conditional on being above the x^{th} percentile of the income distribution, relative to this average minus income at the x^{th} percentile. Applying this formula to the Piketty-Saez data at the 90^{th} percentile of the income distribution implies Pareto coefficients of 2.20 in 1989 and 1.83 in 2014.

Figure 9: Estimated EMG Distribution of Log Household Income (SCF)

income_89-16.pdf

in 2016.

Table 4 reports how this change in inequality changes some key predictions of the model. In this experiment, in order to isolate the impact of changing income dispersion, the mean parameters μ_i are rescaled so that average household income is identical in both years. The table indicates that the observed change in income inequality over this period predicts large changes in graduation rates and net tuition. Given the 1989 distribution, the graduation rate would be 45 percent, and average net tuition would be $\$4,054$. Thus, the model predicts that rising income inequality (all else held equal) reduced the graduation rate by 17.7 percent and increased net tuition by 71.8 percent. This increase is larger than the actual increase in average net tuition observed over this period, indicating that rising income inequality was likely a key factor driving up tuition. Average net tuition rises because households at the top of the income distribution became much richer and thus more willing to pay for expensive high-quality colleges. At the same time, the counterpart to fast income growth at the top of the distribution was declining relative income for households in the middle of the distribution. These households were close to indifferent about attending college in 1989, and these income losses therefore drive down college attendance.

Figure 10 illustrates the changes in the equilibrium quality distribution, in the equilibrium ability mix by quality, and in equilibrium tuition schedules. Increasing income inequality reduces the equilibrium supply of public and low-quality private schools while increasing the number of spots at high-quality private schools. This finding helps to explain why average private net tuition almost triples. Across most of the quality distribution, the share of high-ability students in college declines, indicating that colleges are relying more on instructional spending and less on peer effects to maintain quality. The fact that the aggregate share of students who are high ability is unchanged reflects the compositional shift toward higher-quality (and higher-average-ability) colleges.

One might suspect that rising income inequality would lead to college attendance being driven more by income and less by ability, and thus that increasing inequality would reduce the average ability of college students. However, this does not happen in equilibrium because

Figure 10: Effect of Changing Income Inequality

calibration_distribution_inequality.pdf

Table 4: Changing Income Inequality: 2016 versus 1990

	2016 income dist.			1990 income dist.			Growth Total
	Total	Public	Private	Total	Public	Private	90-16 (%)
Graduation	0.370	0.258	0.113	0.450	0.323	0.127	-17.7
Net tuition, $	6,966	3,788	14,225	4,054	3,515	5,421	71.8
Quality $\frac{q}{\kappa}$	4.76	4.62	5.08	4.45	4.59	4.09	6.9
Share high ability η	0.669	0.715	0.565	0.670	0.727	0.715	-0.2
Expenditure e, $	15,015	11,892	22,147	12,106	11,618	13,343	24.0

increasing inequality also increases the demand for high-ability students. In particular, because the quality production technology features decreasing returns to expenditure, colleges seeking to satisfy increased demand for quality would like to increase both expenditures and average ability, all else equal. Thus, increased demand for high-quality colleges indirectly increases the relative demand for high ability students, which translates into larger institutional tuition discounts for high-ability students (see the bottom panels of Figure 10). This suggests that rising income inequality has played a role in generating the observed growth in institutional financial aid and the associated rising gap between sticker and net tuition (see Figure 1).

In the experiment here, nontuition revenue (conditional on public or private status) is held constant, and thus growth in average net tuition and growth in expenditure per student are closely related, reflecting two sides of the college sectors' balance sheet. Growth in both tuition and expenditure reflects colleges responding to increased demand for quality from households at the top of the income distribution. But how much additional quality does this extra expenditure actually deliver?

Table 4 indicates that our experiment of increasing income inequality increases average college quality by only 6.9 percent, compared to increases in expenditure and net tuition of 24.0 percent and 71.8 percent respectively.[25] Thus, only a small portion of the model growth in net tuition reflects higher average quality. The increase in average quality is small for two reasons. The first is the presence of peer effects: expenditure is only one quality-enhancing

[25] The percentage increase in tuition exceeds the percentage increase in expenditure due to the presence of (constant) subsidies in the former.

input, and the other — average peer ability — does not change in response to rising income inequality. If quality were equal to expenditure, as in a conventional non-club-good model, then the 24.0 percent increase in observed expenditure would correspond to an identical increase in quality.[26]

The second reason the increase in average quality is small is that the increase in expenditure is concentrated in colleges at the top of the quality distribution, where expenditure is already high and the marginal product of additional spending is low. Given equal factor shares ($\theta = 0.5$), our production technology implies that average college quality is given by

$$E[q] = E\left[\bar{a}(q)^{0.5}e(q)^{0.5}\right] = \sqrt{E[\bar{a}] - var\left(\sqrt{\bar{a}}\right)}\sqrt{E[e] - var\left(\sqrt{e}\right)} + cov\left(\sqrt{\bar{a}}, \sqrt{e}\right).$$

When income inequality is increased, the boost to average quality from higher average expenditure is partially offset by a rise in the variance of (the square root of) expenditure, coupled with a fall in the correlation between average ability and expenditure. If we were to observe the same changes in average ability and expenditure in a model with only one type of college (so that $E[q] = \sqrt{E[\bar{a}]}\sqrt{E[e]}$), then average quality would increase by 11.2 percent.[27]

We conclude that in order to properly quantify the rise in average college quality associated with a rise in income inequality, and thus to properly measure growth in quality-adjusted tuition, it is important both to model peer effects and to allow for heterogeneity in college quality: abstracting from either feature leads one to overstate the increase in average quality and thus to understate growth in quality-adjusted tuition.

5.2 Decomposition: 1990 versus 2016

We now use the model to decompose the drivers of observed changes in tuition and graduation rates from 1990 to 2016. Our calibration for 1990 is symmetric to the 2016 calibration described in Section 4. Parameter values for the two years are reported in Table 5. The

[26] We have explicitly solved a non-club-good version of the model, corresponding to $\theta = 0$. In this model, we recalibrated φ and κ to replicate the same graduation and net tuition targets for 2016. Equilibrium expenditure (and quality) increase by 23.0 percent when moving from the 1989 to the 2016 income distribution parameters.

[27] We have not explicitly solved a model with only one college quality. The only model we could come up with in which a degenerate college quality distribution can emerge in equilibrium is one in which all households are identical in terms of their income.

Table 5: Parameters 2016 versus 1990

		2016	1990
Income distribution	$E[y]$	1.00	0.80
	σ^2	0.55	0.48
	α	1.67	2.40
Preferences	φ	0.031	0.022
	κ	0.0332	0.0211
Policies	$s_1 - \phi$	0.0145 ($\$8,111$)	0.0119 ($\$6,680$)
	$s_2 - \phi$	0.0142 ($\$7,930$)	0.0053 ($\$3,007$)
	$\overline{T}$	0.038	0.004
	$\underline{Q}$	0.120	0.149
Technology	p	1.000	0.914

technology parameters θ and $\{a_l, a_h\}$ are assumed time invariant. The income distribution parameters are those estimated from the 1989 wave of the SCF. The preference parameters for 1990 are calibrated to target average net tuition in 1990 ($\$4,865$) and the fraction of young adults with a bachelor's degree (24.2 percent). This implies smaller values for both φ and κ in 1990 relative to 2016. In the next subsection, we translate this change in φ into a change in the college wage premium for the alternative model interpretation in which higher college quality translates into higher child labor earnings.

The estimated subsidies per student in 1990 (net of fixed costs) are smaller than in 2016, especially for private schools. This tends to make public schools relatively more attractive. Replicating relative graduation rates and tuition across the public and private sectors then requires tighter bounds on tuition and quality: a lower tuition cap $\overline{T}$ and a higher quality threshold $\underline{Q}$ shrink the public sector's market share by forcing public schools to operate within a relatively narrow quality range.

We also introduce one more time-varying parameter to the model, which is the price p of the variable instructional input, e. Varying p allows us to assess whether rising faculty salaries play a significant role in accounting for rising tuition (see, e.g., Jones and Yang 2016). We estimate that this price rose by 9.4 percent in real terms over the 1990–2016 period.[28] Table 6 shows the impact of changing various subsets of parameters to their 1990 values, holding all other parameters at their 2016 values. When all parameters are changed

[28]Formally, we assume that z_t units of the consumption good y are required to produce one unit of the instructional input e. Given a competitive market for e, the equilibrium price of e at date t is then $p_t = z_t$. We normalize $z_{2016} = 1$.

Table 6: Decomposition

Calibration		All	Public	Private
2016	*baseline (model = 2016 data)*			
	Enrollment	37.0	25.8	11.3
	Net tuition	6,966	3,788	14,225
(1)	*1989 inequality* σ^2, α			
	Enrollment	45.0	32.3	12.7
	Net tuition	4,054	3,515	5,421
(2)	*1989 mean income* $E[y]$			
	Enrollment	28.8	20.1	8.7
	Net tuition	5,912	3,378	11,786
(3)	*1990 preferences* φ, κ			
	Enrollment	33.1	20.2	12.9
	Net tuition	4,001	3,095	5,416
(4)	*1990 policy: subsidies* s_1, s_2			
	Enrollment	33.3	31.5	1.8
	Net tuition	9,675	5,241	86,080
(5)	*1990 policy: caps* $\overline{T}$, $\underline{Q}$			
	Enrollment	37.0	5.2	31.8
	Net tuition	6,954	2,016	7,764
(6)	*1990 instruction cost* p			
	Enrollment	38.0	28.5	9.5
	Net tuition	6,787	3,271	17,367
1990	*2016 + (1)-(6) (model = 1990 data)*			
	Enrollment	24.2	17.1	7.1
	Net tuition	4,865	2,016	11,781

simultaneously, the model replicates (by construction) the observed 1990 graduation and net tuition values.

We have already discussed experiment (1), which isolates the effect of increasing income inequality. Increasing average income — experiment (2) — raises households' willingness to pay for college and pushes up both the graduation rate and net tuition. The graduation rate rises 8.2 percentage points and net tuition rises 17.8 percent in response to 24.7 percent growth in average income.[29] The rise in net tuition is smaller than the rise in income because the new college entrants choose relatively cheap low-quality colleges.[30]

The next experiment (3) isolates the effect of changes in preference parameters. The increase in φ implies a greater willingness to pay for high-quality colleges and thus a large 74.1 percent increase in average net tuition. At the same time, the graduation rate rises only 3.9 percentage points because the reservation utility parameter κ also increases.

In experiment (4) we reduce subsidies (net of fixed costs) to their 1990 values. Lower subsidies translate into much higher tuition and a smaller graduation rate. The fact that nontuition revenue for private schools was much smaller in 1990 (Table 5) translates into a very small private sector, catering exclusively to very high-quality and high-tuition colleges. Overall, the model implies that growth in non-tuition revenue over time has been an important factor supporting college enrollment, and restraining growth in tuition.[31]

Experiment (5) isolates the impact of estimated changes in the tuition cap and the quality bound. These changes have a negligible impact on both the enrollment rate and average net tuition, but the fact that these constraints were much tighter in 1990 translates into a tiny public sector and a correspondingly expanded private sector. Net tuition and enrollment do not change much because per student net subsidies are very similar in the two sectors in 2016.

Finally, experiment (6) documents the impact of reducing instructional costs to their estimated 1990 value. This reduces average net tuition and raises the graduation rate, but the effects are very small: net tuition falls by less than \$200, and the enrollment rate rises only

[29] If we move both income inequality and average income back to their 1989 values – combining experiments (1) and (2) – then the implied graduation rate is 34.9 percent while tuition is \$3,035.

[30] In this experiment κ, $\overline{T}$ and $\underline{Q}$ are held fixed. If we were to scale these parameters appropriately with growth in income, the rise in net tuition would equal the rise in average income, and the graduation rate would be unchanged (see footnote 17).

[31] We have also conducted experiments in which we increase college subsidies by 10 percent, starting from the baseline 2016 calibration. We find that increasing subsidies to both public and private colleges by \$800 drives down average net tuition by \$1,300 and increases the graduation rate by 2 percentage points.

Table 7: All Structural Changes: 2016 versus 1990

	2016			1990			Growth Total
	Total	Public	Private	Total	Public	Private	90-16 (%)
Net tuition, $	6,966	3,788	14,225	4,865	2,016	11,782	43.2
Quality $\frac{q}{\kappa_{2016}}$	4.76	4.62	5.08	4.52	4.50	4.57	5.3
Share high ability η	0.669	0.715	0.565	0.762	0.943	0.324	-12.2
Expenditure e, $	15,015	11,892	22,147	11,460	8,687	14,778	31.0

1 percentage point. Intuitively, making quality less expensive to produce leads households to increase their chosen college quality but does not change the fraction of income they devote to college very much.

Overall, these experiments highlight the key drivers of observed growth in college tuition in our model. Rising income inequality and a stronger preference for college are the most important factors: each factor can explain more than the entire growth in observed net tuition. Growth in average income is also important, while rising instructional costs play a minimal role. An important force that has worked in the opposite direction and constrained growth in tuition is rising nontuition revenue.

Changes in Tuition versus Changes in Quality: Table 7 summarizes changes in average net tuition, quality, ability, and instructional expenditure when all structural parameters change simultaneously. Recall that in this case, we replicate the observed growth in net tuition by construction.

As in the experiment in which we change only income inequality, we find that the increase in average college quality is much smaller than the increase in average net tuition or the increase in average instructional expenditure.

Changes in Who Goes to College: Belley and Lochner (2007) describe patterns of college attendance by ability (AFQT score) and family income, for college-age individuals in the 1979 and 1997 NLSY. They find that college attendance increases by 13 percentage points between the 1979 and 1997 waves of the survey, which is the same amount of the increase in the graduation rate we find between 1990 and 2016.[32] Almost half of this increase was from students in the bottom half of the AFQT score distribution. Furthermore, these additional

[32] We thank Belley and Lochner for making available the numbers underlying Figure 2 in their paper. Attendance rates for different groups can be computed by combining the information in their Figure 2 with their Table 2. Note that our quantitative exercise focuses on a comparison between 1990 and 2016, while the two NLSY waves offer a comparison of the early 1980s to the early 2000s.

low-ability college students were drawn mainly from the top half of the income distribution: attendance of low-ability students in the top half of the family income distribution almost doubled, rising from 29.7 percent of such students to 54.7 percent, while attendance of low ability students in the bottom half of the income distribution increased much less, from 24.9 percent to 30.8 percent.

Our model broadly reproduces these features of the data. Comparing our calibrations to 1990 and 2016 (Tables 5 and 6), 50.5 percent of the increase in college graduation reflects additional low-ability students attending college, compared to 45.8 percent in the NLSY. All of these extra low-ability students are drawn from the top half of the model family income distribution, compared to 83.8 percent of additional low-ability students in the data. Thus, both model and data are consistent with the message that income has become a more important driver of college attendance relative to student ability.[33]

5.3 College as Consumption versus College as Investment

We now describe an alternative interpretation of our model in which college has an investment component. Here, in addition to the direct consumption value of college, parents also care about future child income, and higher-quality colleges translate into higher child income.[34]

Consider, in particular, the following household utility function:

$$u(c, q, y') = \log c + \beta_1 \log(\kappa + q) + \beta_2 \log(y'), \tag{6}$$

where the preference parameters β_1 and β_2 control, respectively, the strength of the college-

[33]Our decomposition exercise identifies changes in the public college sector as playing a role here. In particular, in 1990 nontuition revenue in the public sector was much larger than in the private sector, but public colleges faced tight tuition and quality caps (Table 5), implying a high equilibrium fraction of high-ability students. By 2016, these constraints were looser, translating into enrollment of more low-ability but full-tuition students.

[34]Carneiro and Heckman (2002) and Hai and Heckman (2017) argue for an important consumption component in college, corresponding to a strong positive correlation between family income and preference for attending college. Dale and Krueger (2014) argue that the financial returns to attending more selective schools might be small, pointing to a large consumption component to attending such schools. In contrast, Belley and Lochner (2007) argue that the college decisions are driven by expected earnings gains and that the consumption value of college is likely negative. Consistent with this view, Zimmerman (2014) documents large earnings gains (relative to tuition costs) for male students just above a state-level GPA cutoff for admission to the Florida State University System using a regression discontinuity approach.

as-consumption and the college-as-investment motives. Child earnings y' are given by

$$y' = (\kappa + q)^{\zeta} a^{\lambda}, \tag{7}$$

where the parameters ζ and λ determine, respectively, the elasticity of future child earnings with respect to college quality and own ability.

Note that this model implicitly assumes that investment in own-child education is the only way to transfer resources across generations. Pareto improvements could be realized by introducing intergenerational credit markets, allowing rich low-ability households to lend to poor high-ability ones. In fact, any efficient allocation would feature a positive relation between ability and college quality, given the complementarity between the two inputs in the earnings production function.

By substituting eq. (6) with eq. (7), it is clear that this model has a reduced form that is identical to the original model (see eq. 1). In particular, given a value for ζ, the return-to-college parameter, any pair (β_1, β_2) satisfying $\beta_1 + \beta_2\zeta = \varphi$ generates identical predictions for all the observables we have discussed so far, including tuition and enrollment. Thus, these observables cannot be used to differentiate between the consumption versus investment motives for attending college.

Where the two motives differ is in their implications for the responses of college tuition and enrollment to *changes* in the college premium. To see this, assume that the preference parameters β_1 and β_2 are time invariant and that the return-to-ability parameter λ is also constant. Now consider an increase over time in the return-to-college parameter ζ. The larger is β_2 (the stronger the college-as-investment motive), the larger will be the equilibrium increases in enrollment and tuition. In contrast, if $\beta_2 = 0$ (college is a pure consumption good), then enrollment and tuition should not respond to changes in the return to college.

We can therefore use evidence on changes in the college premium over our sample period to identify the relative importance of the college-as-consumption versus college-as-investment motives. We start by assuming that the return to ability parameter λ is time invariant. Hendricks and Schoellman (2014) estimate that a one standard deviation in ability raises log wages by 10.4 percentage points. Given that we assume half the population is high ability and half is low ability, the ratio a_h/a_l is two standard deviations, and thus Hendricks and Schoellman's estimate suggests that, given the same quality college, high-ability students should earn 20.8 percent more than their low-ability counterparts; that is, $1.208 = \left(\frac{a_h}{a_l}\right)^{\lambda}$, which implies that $\lambda = 0.248$.

The remaining four parameters, β_1, β_2, ζ_{1990}, and ζ_{2016}, are chosen to replicate four targets. First, the two return to college parameters ζ_{1990} and ζ_{2016} are then chosen so that the simulated model replicates the observed college wage premia in 1990 and 2016. We measure the college premium using Current Population Survey (CPS) data, defining the premium as the ratio of average annual wage income for workers with a bachelor's degree relative to those who have only completed high school.[35] This premium was 1.61 in 1990 and 1.83 in 2016. Note that in both the model and data, the measured college premium partly reflects selection by ability into college. The preference parameters β_1 and β_2 are then set to satisfy $\varphi_t = \beta_1 + \beta_2 \zeta_t$ for both $t = 1990$ and $t = 2016$, so that this extended model generates the same tuition and enrollment patterns as the baseline model described earlier.

This procedure delivers $\zeta_{1990} = 0.1983$, $\zeta_{2016} = 0.3230$, $\beta_1 = 0.0077$, and $\beta_2 = 0.0715$. Thus, the model suggests that college has both a consumption component and an investment component. Defining the consumption share of the total value of college as $\beta_1/(\beta_1 + \beta_2 \zeta_t)$, the consumption share was 35 percent in 1990. By 2016, this share had declined to 25 percent, reflecting a larger college premium.

6 Conclusion

A satisfactory model of the college market is essential for understanding what sorts of potential students go to college, what sorts of colleges they attend, and how much they pay. It is also important for understanding how these features of the college landscape have changed over time, and for exploring the impact of possible policy interventions.

We have developed a competitive model of the college market in which college quality depends on the average ability of attending students. A novel feature of the model is a continuous distribution of college quality. A calibrated version of the model generates a similar distribution of college tuition across colleges to that observed empirically, and also replicates observed positive correlations between tuition on the one hand and measures of student ability and family income on the other.

[35] We restrict the sample to full-time workers aged 40-45. In 1990 college is defined as having completed four years of college and high school as having completed 12th grade.

When we use the model to predict the impact of the rise in top-tail income inequality since 1990, we find that greater income inequality can explain the entire observed increase in average college tuition. By itself, greater income inequality would depress graduation rates, but we find that more generous college subsidies and stronger preference for quality are forces that have pushed more people into college.

We have argued that a perfectly competitive model in which colleges minimize cost offers a reasonable positive theory of observed outcomes in the college market. One could allow colleges to earn rents (and thereby consider alternative objectives to profit maximization) by endowing colleges with an idiosyncratic nonreproducible attribute, such as location or brand name, and endowing households with heterogeneous preferences over this attribute. Another possible extension to the model would be to model heterogeneity in college information about student ability. At the time of admission, students would then be unable to perfectly anticipate the terms of admission offers and would therefore want to apply to multiple schools (see, e.g., Fu 2014). Yet another interesting extension would be to explore the nature of optimal college subsidies, tying subsidies to student attributes (income or ability) or to the quality of the college the student attends, or both (see Findeisen and Sachs 2018). The nature of the optimal intervention will depend on the planner's social welfare function, in addition to whether the laissez-faire allocation is efficient (as when college is a pure consumption good) or inefficient.

One interpretation of the fact that graduation rates have increased during a period of rising income inequality is that college has an investment component, which has increased in importance with a widening college premium. Under this college-as-investment interpretation, the relative roles of student ability versus parental income as drivers of college quality choices will depend on students' ability to borrow to pay for college. Embedding our model of the college market into a life-cycle framework with a quantitative model of the student loan market is one possible avenue for future research in this area (see, for example, Abbott et al. 2016).

Another possible application is to develop a multigenerational extension of the model outlined in Section 5.3 to explore the propagation of inequality across generations. Consider an increase in the financial return to college quality. This will lead to an increase in investment in quality by higher-income households, which will amplify the effect on income inequality in the next generation. In turn, a fatter right tail in the income distribution for that generation will further amplify inequality in college investment. Over successive generations, a small increase in the return to college quality can potentially generate both a large increase in income inequality and a decline in intergenerational mobility.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2016). Education Policy and Intergenerational Transfers in Equilibrium. *Working Paper*.

Belley, P. and L. Lochner (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital 1*(1), 37–89.

Carneiro, P. and J. J. Heckman (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *Economic Journal 112*(482), 705–734.

Caucutt, E. M. (2001). Peer Group Effects in Applied General Equilibrium. *Economic Theory 17*(1), 25–51.

Caucutt, E. M. (2002). Educational Vouchers When There Are Peer Group Effects——Size Matters. *International Economic Review 43*(1), 195–222.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2017). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility. Working Paper 23618, National Bureau of Economic Research.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Cole, H. L. and E. C. Prescott (1997). Valuation Equilibrium with Clubs. *Journal of Economic Theory 74*(1), 19–39.

College Board (2016a). *Trends in College Pricing 2016*. Trends in Higher Education Series. Available at https://trends.collegeboard.org.

College Board (2016b). *Trends in Student Aid 2016*. Trends in Higher Education Series. Available at https://trends.collegeboard.org.

Dale, S. B. and A. B. Krueger (2014). Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data. *Journal of Human Resources 49*(2), 323–358.

Ellickson, B., B. Grodal, S. Scotchmer, and W. R. Zame (1999). Clubs and the Market. *Econometrica 67*(5), 1185–1217.

Epple, D., R. Romano, S. Sarpca, and H. Sieg (2017). A General Equilibrium Analysis of State and Private Colleges and Access to Higher Education in the U.S. *Journal of Public Economics 155*(C), 164–178.

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

Epple, D. and R. E. Romano (1998). Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review 88*(1), 33–62.

Fillmore, I. (2016). Price Discrimination and Public Policy in the U.S. College Market. Working paper, Washington University.

Findeisen, S. and D. Sachs (2018). Optimal Need-based Financial Aid. *Working Paper*.

Fu, C. (2014). Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market. *Journal of Political Economy 122*(2), 225–281.

Golden, D. (2006). The Price of Admission: How America's Ruling Class Buys Its Way into Elite Colleges—and Who Gets Left Outside the Gates. *New York: Three Rivers Press.*.

Gordon, G. and A. Hedlund (2017). Accounting for the Rise in College Tuition. Working Paper 21967, National Bureau of Economic Research.

Hai, R. and J. J. Heckman (2017). Inequality in Human Capital and Endogenous Credit Constraints. *Review of Economic Dynamics 25*, 4–36. Special Issue on Human Capital and Inequality.

Heathcote, J. and H. Tsujiyama (2017). Optimal Income Taxation: Mirrlees Meets Ramsey. Working paper, Federal Reserve Bank of Minneapolis.

Hendricks, L. and T. Schoellman (2014). Student Abilities During the Expansion of US Education. *Journal of Monetary Economics 63*, 19–36.

Jones, J. B. and F. Yang (2016). Skill-Biased Technical Change and the Cost of Higher Education. *Journal of Labor Economics 34*(3), 621–662.

Piketty, T. and E. Saez (2003). Income Inequality in the United States, 1913–1998. *Quarterly Journal of Economics 118*(1), 1–41.

Rothschild, M. and L. J. White (1995). The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs. *Journal of Political Economy 103*(3), 573–586.

Schwartz, A. E. and B. Scafidi (2004). What's Happened to the Price of College? Quality-Adjusted Net Price Indexes for Four-Year Colleges. *Journal of Human Resources 39*(3), 723–745.

Scotchmer, S. (1997). On Price-Taking Equilibria in Club Economies with Nonanonymous Crowding. *Journal of Public Economics 65*(1), 75–88.

Zimmerman, S. D. (2014). The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics 32*(4), 711–754.

7 Not for Publication Appendices

7.1 Theoretical Appendix

Proof of Proposition 1

In this proof we show that for any equilibrium with a potentially nonlinear tuition schedule, we can construct an alternative linear tuition schedule that is consistent with the equilibrium allocation. Thus, any equilibrium can be supported by a tuition schedule that is linear with respect to ability. The proof is simple if there are no missing markets (*i.e.*, if $\eta^i(q) > 0$, $\forall i$). In contrast, when only a subset of ability types attend colleges of a given quality, there may be a set of tuition schedules consistent with the same allocation. We show that there must be a linear tuition schedule within this set.

Proof.

Denote the set of active college markets of quality q by $A^+(q) = \{a_i : \eta^i(q) > 0\}$. This set is nonempty given that $\chi(q) > 0$. Let $a_{\max}$ and $a_{\min}$ be the maximum and minimum elements in this set.

Case 1: $A^+(q)$ is not a singleton set (so that $a_{\max} > a_{\min}$):

In this case, define

$$d(q) = -\frac{t(q, a_{\max}) - t(q, a_{\min})}{a_{\max} - a_{\min}}$$
$$b(q) = t(q, a_{\min}) + d(q)(a_{\min} - a_1).$$

Now we claim that for any a_i,

$$t(q, a_i) = b(q) - d(q)(a_i - a_1).$$

We prove this by contradiction. Suppose that for some j

$$t(q, a_j) > b(q) - d(q)(a_j - a_1).$$

Then one can show that a college can increase profit by shifting students to a_j from either $a_{\min}$ or $a_{\max}$ to replicate the desired ability level in college and can continue doing so until either $\eta^{\min}$ or $\eta^{\max}$ is zero. But this is a contradiction to the assumption that both $a_{\max}$ and $a_{\min}$ belong to the set of active markets $A^+(q)$. Thus, it must be that

$$t(q, a_j) \leq b(q) - d(q)(a_j - a_1) \text{ for any } j.$$

Now suppose that $t(q, a_j) < b(q) + d(q)(a_j - a_1)$ for some j. It must then be that $\eta^j(q) = 0$ in equilibrium. Otherwise, the college could shift admissions from a^j students to other ability levels, maintaining the desired average ability and making greater profit. Thus, in equilibrium it must be the case that both the supply $s(q, a_j)$ and the demand $d(q, a_j)$ in this particular market (q, a_j) are zero. Now replace the tuition value $t(q, a_j)$ with $\tilde{t}(q, a_j)$ defined by

$$\tilde{t}(q, a_j) = b(q) - d(q)(a_j - a_1).$$

Note that at the new level of tuition $\tilde{t}(q, a_j)$, college demand for students $d(q, a_j)$ will still be zero (because colleges are indifferent between admitting students with ability a_j or a group of students with average ability a_j). And the supply of students is zero as well because it is now more costly for the households to pick this college $\tilde{t}(q, a_j) > t(q, a_j)$. Thus the market still clears under the new level of tuition $\tilde{t}(q, a_j)$. Thus, without loss of generality we can treat $\tilde{t}(q, a_j)$ as the equilibrium tuition value. And we have finished the first part of the proof.

Case 2: $A^+(q) = \{a_i : \eta^i(q) > 0\}$ is a singleton set:

In this case, first let a_m be the unique element of $A^+(q)$. Define a set of discount rates $D^<(q) = \left\{d_i(q), i < m : d_i(q) = -\frac{t(q,a_i)-t(q,a_m)}{a_i-a_m}\right\}$, which is the set of tuition slopes between a_m and other ability levels lower than a_m. If this set of nonempty, pick the greatest element in this set $D^<(q)$, $d_n(q)$, and denote the associated a as a_n. Define

$$b_n(q) = t(q, a_m).$$

Now we claim that for any a_j, it must be the case that

$$t(q, a_j) \leq b_n(q) - d_n(q)(a_j - a_m).$$

To see this, note that for any $a_j < a_m$, by definition the slope $d_j(q) \leq d_n(q)$, and thus

$$\begin{aligned} t(q, a_j) &= t(q, a_m) - d_j(q)(a_j - a_m) \\ &= b_n(q) - d_j(q)(a_j - a_m) \\ &\geq b_n(q) - d_n(q)(a_j - a_m), \end{aligned}$$

where the last inequality holds because $a_j - a_m < 0$.

Next, for any $a_j > a_m$, we can show this by contradiction. Suppose that

$$t(q, a_j) > b_n(q) - d_n(q)(a_j - a_m).$$

Then the college can use a mix of a_j and a_n students to replicate ability level a_m (since $a_j > a_m > a_n$, such a mix is feasible). This yields greater profit for the college. But this is a contradiction to optimality. Thus, we prove that for any a_j

$$t(q, a_j) \leq b_n(q) - d_n(q)(a_j - a_m).$$

And similarly to the first part of the proof, we can replace $t(q, a_j)$ with

$$\tilde{t}(q, a_j) = b_n(q) - d_n(q)(a_j - a_m)$$

and maintain market clearing. The very last step is to show that even when the set $D^<(q)$ is empty, we still have a linear tuition schedule. To see this, define

$$D^>(q) = \left\{d_i(q), i > m : d_i(q) = -\frac{t(q, a_i) - t(q, a_m)}{a_i - a_m}\right\}.$$

This set must be nonempty given that $D^<(q)$ is empty. Pick the smallest element in this set and denote it $d_l(q)$ with associated ability level $a_l(q)$. Also define

$$b_l(q) = t(q, a_m).$$

Now we claim that for any a_j,

$$t(q, a_j) \leq b_l(q) - d_l(q)(a_j - a_m)$$

The alternative case in which this inequality is not satisfied would violate the fact that $d_l(q)$ is the smallest element of $D^{>}(q)$ and that $D^{<}(q)$ is empty. Thus, again similarly to the first part of the proof, we can replace $t(q, a_j)$ with

$$\tilde{t}(q, a_j) = b_n(q) - d_n(q)(a_j - a_m)$$

and all the market-clearing conditions are satisfied. This completes the proof. ■

Proof of Proposition 2

We first define a feasible allocation in the benchmark economy.

Definition. *A feasible allocation is a set of functions* $\chi(q)$, $c^i(y)$, $q^i(y)$, $\eta^i(q)$, $e(q)$ *such that*:

1. The goods market clears:
$$\sum_{i\leq I}\mu_i\int_0^\infty c^i(y)dF_i(y)+\int_0^\infty e(q)d\chi(q)+(1-\chi(0))(\omega+\phi)=\sum_{i\leq I}\mu_i\int_0^\infty ydF_i(y).$$

2. The college markets clear. For all i and $Q\subset\mathbb{R}^+$,
$$\mu_i\int \mathbb{1}_{\{q^i(y)\in Q\}}dF_i(y) \quad = \quad \int_Q \eta^i(q)d\chi(q),$$
where $\mathbb{1}_{\{.\}}$ is an indicator function.

Definition. *An allocation* $\chi(q)$, $c^i(y)$, $q^i(y)$, $\eta^i(q)$, $e(q)$ *is Pareto optimal if 1) it is a feasible allocation and 2) there does not exist a feasible allocation* $\chi'(q)$, $c^{i\prime}(y)$, $q^{i\prime}(y)$, $\eta^{i\prime}(q)$, $e'(q)$ *such that* $u\left(c^{i\prime}\left(y\right),q^{i\prime}\left(y\right)\right)\geq u\left(c^{i}\left(y\right),q^{i}\left(y\right)\right)$ *for almost every* (i,y) and $u\left(c^{i\prime}\left(y\right),q^{i\prime}\left(y\right)\right)>u\left(c^{i}\left(y\right),q^{i}\left(y\right)\right)$ *for some set* (i,y) *of positive measure.*

We now prove the First Welfare Theorem. The proof here closely mirrors the standard proof of the Welfare Theorem.

Theorem. *Assume that* u *exhibits local nonsatiation. If* $\chi(q)$, $c^i(y)$, $q^i(y)$, $\eta^i(q)$, $e(q)$ *is a competitive equilibrium allocation, then it is Pareto efficient.*

Proof. Suppose to the contrary that there exists an alternative feasible allocation $\chi'(q)$, $c^{i\prime}(y)$, $q^{i\prime}(y)$, $\eta^{i\prime}(q)$, $e'(q)$ such that $u\left(c^{i\prime}\left(y\right),q^{i\prime}\left(y\right)\right)\geq u\left(c^{i}\left(y\right),q^{i}\left(y\right)\right)$ for almost every (i,y) and $u\left(c^{i\prime}\left(y\right),q^{i\prime}\left(y\right)\right)>u\left(c^{i}\left(y\right),q^{i}\left(y\right)\right)$ for some positive measure set of (i,y). Denote $t\left(q,a_i\right)$ the equilibrium tuition function associated with the competitive equilibrium allocation. Then, from the local nonsatiation assumption, we know that the alternative allocation must lie outside households' budget set:
$$c^{i\prime}\left(y\right)+t\left(q^{i\prime}\left(y\right),a_i\right)\geq y-\omega\mathbb{1}_{\{q^{i\prime}(y)>0\}}\text{ for almost every }\left(i,y\right).$$

Otherwise, $c^{i}\left(y\right),q^{i}\left(y\right)$ would not be individually rational given the tuition functions. In addition,
$$c^{i\prime}\left(y\right)+t\left(q^{i\prime}\left(y\right),a_i\right)>y-\omega\mathbb{1}_{\{q^{i\prime}(y)>0\}}\text{ for some positive measure set.}$$

Summing up the above equations across households of different abilities and income, we get
$$\sum_{i\leq I}\mu_i\int c^{i'}\left(y\right)dF_i\left(y\right)+\sum_{i\leq I}\mu_i\int t\left(q^{i\prime}\left(y\right),a_i\right)dF_i\left(y\right)>\sum_{i\leq I}\mu_i\int ydF_i\left(y\right)-(1-\chi'(0))\omega. \quad (8)$$

Note that under the alternative feasible allocation, aggregate enrollment is given by $1-\chi'(0)$.

Now turn to the college sector. Since the equilibrium allocation maximizes the colleges' profit under the competitive price vector, the alternative allocation must be (weakly) inferior to the competitive equilibrium allocation under the competitive tuition schedule. Thus, the colleges must make nonpositive profit:

$$\sum_{i\leq I}\eta^{i\prime}(q)t(q,a_i)-e'(q)-\phi\leq 0.$$

Since this equation holds for all colleges, the aggregate profit made by the college sector must be nonpositive, which in turn implies that the aggregate tuition revenue is no greater than the college expenditure:

$$\sum_{i\leq I}\int t\left(q^{i\prime}(y),a_i\right)dF_i(y)\leq\int_0^\infty e'(q)d\chi(q)+(1-\chi'(0))\phi. \tag{9}$$

But eqs. 8 and 9 together imply that the aggregate resources constraint is violated under the alternative allocation:

$$\begin{aligned}
& \sum_{i\leq I}\mu_i\int_0^\infty c^{i\prime}(y)dF_i(y)+\int_0^\infty e'(q)d\chi(q)+(1-\chi'(0))(\omega+\phi)\\
\geq\ & \sum_{i\leq I}\mu_i\int_0^\infty c^{i\prime}(y)dF_i(y)+\sum_{i\leq I}\int t\left(q^{i\prime}(y),a_i\right)dF_i(y)+(1-\chi'(0))\omega\\
>\ & \sum_{i\leq I}\int ydF_i(y),
\end{aligned}$$

where the first weak inequality follows from 9 and the second strict inequality follows from 8. Thus we have a contradiction. This establishes the Pareto-optimality of the competitive allocation.

Proof of Proposition 3

Theorem. Suppose education is a pure club good ($\theta = 1$) and household utility is given by

$$\log(c) + \log(\kappa + q),$$

and the income distribution for both high and low ability is uniform in some interval $\left[\mu_y - \frac{1}{2}\Delta_y, \mu_y + \frac{1}{2}\Delta_y\right]$. Then the college distribution is given by

$$\begin{aligned}
\chi(Q) &= \frac{2}{a_h - a_l}\frac{2}{4+\pi}\int_Q \left[(1-\eta(q))^2 + \eta(q)^2\right]^{-2} dq \;\; \forall Q \subset (a_l, a_h) \\
\chi(a_h) &= \chi(a_l) = \frac{2}{4+\pi},
\end{aligned}$$

and the tuition function is given by

$$t^i(q) = \mu_y \frac{q - a_i}{\kappa + q}\left[1 - \frac{2}{4+\pi}\frac{\Delta_y}{\mu_y}\arctan(1 - 2\eta(q))\right], i = h, l.$$

Proof. As a first step, the household problem given income y is

$$\max_{c,q} \log(c) + \log(\kappa + q)$$

$$c + t^i(q) \leq y.$$

Assuming that $t^i(q)$ is differentiable (verified later), the first-order condition for households is

$$t^{i\prime}(q) = \frac{y - t^i(q)}{\kappa + q}.$$

Denoting by $y(q)$ the income of the households attending colleges of quality q in equilibrium, we have

$$\begin{aligned}
t^{i\prime}(q) &= \frac{y(q) - t^i(q)}{\kappa + q} \\
&= -\frac{1}{\kappa + q}t^i(q) + \frac{y^i(q)}{\kappa + q}.
\end{aligned}$$

This is a linear ODE that can be solved using the integrating factor method. Define the integrating factor for low ability tuition $v^l(q)$ as

$$\begin{aligned}
v^l(q) &= \int_{a_l}^{q} \frac{1}{\kappa + q'} dq' \\
&= \log\frac{\kappa + q}{\kappa + a_l}.
\end{aligned}$$

Thus,

$$\exp\left(v^l(q)\right) = \frac{\kappa + q}{\kappa + a_l}$$

$$\begin{aligned}
\exp\left(v^{l}(q)\right)t^{a\prime}(q)+\frac{1}{\kappa+q}\exp\left(v^{l}(q)\right)t^{a}(q) &= \exp\left(v^{l}(q)\right)\frac{y(q)}{\kappa+q} \\
\left[\exp\left(v^{l}(q)\right)t^{a}(q)\right]' &= \exp\left(v^{l}(q)\right)\frac{y(q)}{\kappa+q} \\
\int_{a_l}^{q}\left[\exp\left(v^{l}(q)\right)t^{a}(q)\right]' dq &= \int_{a_l}^{q}\exp\left(v^{l}(q)\right)\frac{y(q)}{\kappa+q}dq \\
\exp\left(v^{l}(q)\right)t^{l}(q)-\exp(a_l)\,t^{l}(a_l) &= \int_{a_l}^{q}\exp\left(v^{l}(q)\right)\frac{y(q)}{\kappa+q}dq \\
&= \int_{a_l}^{q}\frac{y(q)}{\kappa+a_l}dq.
\end{aligned}$$

We know from the zero profit condition that

$$t^{l}(a_l)=0$$

$$\begin{aligned}
\exp\left(v^{l}(q)\right)t^{l}(q) &= \int_{a_l}^{q}\frac{y(q')}{\kappa+a_l}dq' \\
t^{l}(q) &= \exp\left(-v^{l}(q)\right)\int_{a_l}^{q}\frac{y(q')}{\kappa+a_l}dq' \\
&= \frac{\kappa+a_l}{\kappa+q}\int_{a_l}^{q}\frac{y(q')}{\kappa+a_l}dq' \\
&= \int_{a_l}^{q}\frac{y^{l}(q')}{\kappa+q}dq'.
\end{aligned}$$

The integrating factor for the high-ability type is given by

$$\begin{aligned}
v^{h}(q) &= \int_{a_h}^{q}\frac{1}{\kappa+q'}dq' \\
&= \log\frac{\kappa+q}{\kappa+a_h}.
\end{aligned}$$

We can follow the same procedure and obtain an expression for the high-ability tuition function:

$$\exp\left(v^{h}(q)\right)t^{h}(q)-\exp\left(v^{h}(a_h)\right)t^{h}(a_h)=\int_{a_h}^{q}\exp\left(v^{h}(q)\right)\frac{y(q)}{\kappa+q}dq.$$

The zero profit condition for the $q=a_h$ college implies

$$t^{h}(a_h)=0.$$

Thus,

$$\begin{aligned}
t^{h}(q) &= \int_{a_h}^{q}\frac{y(q)}{\kappa+q}dq \\
&= -\int_{q}^{a_h}\frac{y^{h}(q)}{\kappa+q}dq.
\end{aligned}$$

Now we derive the income function $y^i(q)$, given uniformly distributed income and any college distribution function $\chi(q)$:

$$\begin{aligned} y^h(q) &= \mu_y + \frac{1}{2}\Delta_y - \Delta_y \int_q^{a_h} \chi(q') \frac{q' - a_l}{a_h - a_l} dq' \\ y^l(q) &= \mu_y + \frac{1}{2}\Delta_y - \Delta_y \int_q^{a_h} \chi(q') \frac{a_h - q'}{a_h - a_l} dq'. \end{aligned}$$

Now, we would like to solve for $\chi(q)$ from the zero profit condition $\pi(q) = 0$. We conjecture that there is a strictly positive measure of high-ability students going to colleges of quality $q = a_h$. Denote that mass $\chi(a_h)$. Thus,

$$y^h(q) = \mu_y + \frac{1}{2}\Delta_y - \Delta_y \left(\int_q^{a_h} \chi(q') \frac{q' - a_l}{a_h - a_l} dq' + \chi(a_h) \right).$$

Write out the expression for $\pi(q)$:

$$\begin{aligned} \pi(q) &= \frac{q - a_l}{a_h - a_l} t^h(q) + \frac{a_h - q}{a_h - a_l} t^l(q) \\ &= -\frac{q - a_l}{a_h - a_l} \int_q^{a_h} \frac{y^h(q')}{\kappa + q} dq + \frac{a_h - q}{a_h - a_l} \int_{a_l}^{q} \frac{y^l(q')}{\kappa + q} dq' \\ &= 0. \end{aligned}$$

Cancelling out some terms, we have that for any q

$$(a_h - q) \int_{a_l}^{q} y^l(q')\, dq' - (q - a_l) \int_q^{a_h} y^h(q')\, dq' = 0$$

Substitute in expressions for $y^l(q)$ and $y^h(q)$:

$$(a_h - q) \left(\int_{a_l}^{q} \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \int_{q'}^{a_h} \chi(x) \frac{a_h - x}{a_h - a_l} dx \right) dq' \right) -$$

$$(q - a_l) \left(\int_q^{a_h} \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \left(\int_{q'}^{a_h} \chi(x) \frac{x - a_l}{a_h - a_l} dx + \chi(a_h) \right) \right) dq' \right) = 0.$$

Differentiate with respect to q :

$$\begin{aligned} &- \left(\int_{a_l}^{q} \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \int_x^{a_h} \chi(x) \frac{a_h - x}{a_h - a_l} dx \right) dq' \right) \\ &+ (a_h - q) \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \int_q^{a_h} \chi(x) \frac{a_h - x}{a_h - a_l} dx \right) \\ &- \left(\int_q^{a_h} \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \left(\int_x^{a_h} \chi(x) \frac{x - a_l}{a_h - a_l} dx + \chi(a_h) \right) \right) dq \right) \\ &+ (q - a_l) \left(\mu_y + \frac{1}{2}\Delta_y - \Delta_y \left(\int_q^{a_h} \chi(x) \frac{x - a_l}{a_h - a_l} dx + \chi(a_h) \right) \right) = 0. \end{aligned}$$

Differentiate again with respect to q :

$$\begin{aligned}
&-\left(\mu_y+\frac{1}{2}\Delta_y-\Delta_y\int_q^{a_h}\chi\left(x\right)\frac{a_h-x}{a_h-a_l}dx\right)\\
&-\left(\mu_y+\frac{1}{2}\Delta_y-\Delta_y\int_q^{a_h}\chi\left(x\right)\frac{a_h-x}{a_h-a_l}dx\right)\\
&+\left(a_h-q\right)\Delta_y\chi\left(q\right)\frac{a_h-q}{a_h-a_l}+\left(\mu_y+\frac{1}{2}\Delta_y-\Delta_y\left(\int_x^{a_h}\chi\left(x\right)\frac{x-a_l}{a_h-a_l}dx+\chi\left(a_h\right)\right)\right)\\
&+\left(\mu_y+\frac{1}{2}\Delta_y-\Delta_y\left(\int_q^{a_h}\chi\left(x\right)\frac{x-a_l}{a_h-a_l}dx+\chi\left(a_h\right)\right)\right)+\left(q-a_l\right)\Delta_y\left(\chi\left(q\right)\frac{q-a_l}{a_h-a_l}\right)=0.
\end{aligned}$$

Collect terms:

$$\begin{aligned}
&-2\left(-\Delta_y\int_q^{a_h}\chi\left(x\right)\frac{a_h-x}{a_h-a_l}dx\right)+\left(a_h-q\right)\Delta_y\chi\left(q\right)\frac{a_h-q}{a_h-a_l}\\
&2\left(-\Delta_y\left(\int_q^{a_h}\chi\left(x\right)\frac{x-a_l}{a_h-a_l}dx+\chi\left(a_h\right)\right)\right)+\left(q-a_l\right)\Delta_y\left(\chi\left(q\right)\frac{q-a_l}{a_h-a_l}\right)=0.
\end{aligned}$$

Note that Δ_y can be factored out, and we arrive at a functional equation $\chi\left(q\right)$ that is independent of the income distribution parameters:

$$\begin{aligned}
&2\int_q^{a_h}\chi\left(x\right)\left(a_h-x\right)dx+\left(a_h-q\right)^2\chi\left(q\right)\\
&-2\left(\int_q^{a_h}\chi\left(x\right)\left(x-a_l\right)dx+\chi\left(a_h\right)\left[a_h-a_l\right]\right)+\left(q-a_l\right)^2\chi\left(q\right)\\
=\;&0
\end{aligned}$$

Thus, we have the following integral equation:

$$\left[\left(a_h-q\right)^2+\left(q-a_l\right)^2\right]\chi\left(q\right)+2\int_q^{a_h}\chi\left(x\right)\left(a_h+a_l-2x\right)dx=2\chi\left(a_h\right)\left[a_h-a_l\right]$$

$$\chi\left(q\right)=\frac{-2}{\left[\left(a_h-q\right)^2+\left(q-a_l\right)^2\right]}\int_{a_h}^{q}\chi\left(x\right)\left(2x-a_h-a_l\right)dx+\frac{2\chi\left(a_h\right)\left[a_h-a_l\right]}{\left(a_h-q\right)^2+\left(q-a_l\right)^2}.$$

This is a Volterra equation of the second type with degenerate kernels, which happens to have an analytical solution. Define the following objects:

$$\begin{aligned}
g\left(q\right)&=\frac{-2}{\left[\left(a_h-q\right)^2+\left(q-a_l\right)^2\right]}\\
h\left(x\right)&=\left(2x-a_h-a_l\right)\\
f\left(q\right)&=\frac{2\chi\left(a_h\right)\left[a_h-a_l\right]}{\left(a_h-q\right)^2+\left(q-a_l\right)^2}.
\end{aligned}$$

and we have

$$\chi\left(q\right)=f\left(q\right)+\int_{a_h}^{q}R\left(q,x\right)f\left(x\right)dx$$

where

$$\begin{aligned}
R(q,x) &= g(q)\,h(x)\exp\left[\int_x^q g(s)\,h(s)\,ds\right] \\
&= \frac{-2(2x-a_h-a_l)}{\left[(a_h-q)^2+(q-a_l)^2\right]}\exp\left[\int_x^q \frac{-2(2s-a_h-a_l)}{\left[(a_h-s)^2+(s-a_l)^2\right]}ds\right] \\
&= \frac{-2(2x-a_h-a_l)}{\left[(a_h-q)^2+(q-a_l)^2\right]}\exp\left[-\int_x^q \frac{1}{\left[(a_h-s)^2+(s-a_l)^2\right]}d\left[(a_h-s)^2+(s-a_l)^2\right]\right] \\
&= \frac{-2(2x-a_h-a_l)}{\left[(a_h-q)^2+(q-a_l)^2\right]}\exp\left[-\log\frac{(a_h-q)^2+(q-a_l)^2}{(a_h-x)^2+(x-a_l)^2}\right] \\
&= \frac{-2(2x-a_h-a_l)\left((a_h-x)^2+(x-a_l)^2\right)}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\int_{a_h}^q R(q,x)\,f(x)\,dx \\
&= \int_{a_h}^q \frac{-2(2x-a_h-a_l)\left((a_h-x)^2+(x-a_l)^2\right)}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}\frac{2\chi(a_h)[a_h-a_l]}{(a_h-x)^2+(x-a_l)^2}dx \\
&= \frac{-2\chi(a_h)[a_h-a_l]}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}\left[(a_h-q)^2+(q-a_l)^2\right]|_{a_h}^q \\
&= \frac{-2\chi(a_h)[a_h-a_l]}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}\left[(a_h-q)^2+(q-a_l)^2-(a_h-a_l)^2\right].
\end{aligned}$$

Thus,

$$\begin{aligned}
\chi(q) &= \frac{2\chi(a_h)[a_h-a_l]}{(a_h-q)^2+(q-a_l)^2}+\frac{-2\chi(a_h)[a_h-a_l]}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}\left[(a_h-q)^2+(q-a_l)^2-(a_h-a_l)^2\right] \\
&= \frac{2\chi(a_h)[a_h-a_l]^3}{\left[(a_h-q)^2+(q-a_l)^2\right]^2} \\
&= M(q)\,\chi(a_h),
\end{aligned}$$

where

$$M(q)=\frac{2[a_h-a_l]^3}{\left[(a_h-q)^2+(q-a_l)^2\right]^2}.$$

Thus, we can derive the value of $\chi(a_h)$ from

$$\begin{aligned}
\int_{a_l}^{a_h}\chi(q)\,\frac{q-a_l}{a_h-a_l}dq+\chi(a_h) &= 1 \\
\chi(a_h) &= \frac{1}{1+\int_{a_l}^{a_h}M(q)\,\frac{q-a_l}{a_h-a_l}dq}.
\end{aligned}$$

Now,

$$\begin{aligned}
& \int_{a_l}^{a_h} M(q) \frac{q-a_l}{a_h-a_l} dq \\
= \quad & 2[a_h-a_l]^2 \int_{a_l}^{a_h} \frac{(q-a_l)}{\left[(a_h-q)^2+(q-a_l)^2\right]^2} dq \\
= \quad & 2[a_h-a_l]^2 \frac{\frac{(a_h-a_l)(q-a_h)}{(a_h-q)^2+(q-a_l)^2} + \arctan\frac{2q-a_l-a_h}{a_h-a_l}}{2(a_h-a_l)^2} |_{a_l}^{a_h} \\
= \quad & \frac{(a_h-a_l)(q-a_h)}{(a_h-q)^2+(q-a_l)^2} + \arctan\frac{2q-a_l-a_h}{a_h-a_l} |_{a_l}^{a_h} \\
= \quad & 0 + \arctan(1) + 1 - \arctan(-1) \\
= \quad & 1 + 2\arctan(1) \\
= \quad & 1 + \frac{\pi}{2}.
\end{aligned}$$

Thus,

$$\chi(a_h) = \frac{1}{2+\frac{\pi}{2}}$$

$$\chi(q) = \frac{2[a_h-a_l]^3}{\left[(a_h-q)^2+(q-a_l)^2\right]^2} \frac{1}{2+\frac{\pi}{2}}.$$

Now we know that all low-ability types not going to $q > a_l$ will go to $q = a_l$ college. That is given by

$$\begin{aligned}
\chi(a_l) &= 1 - \int_{a_l}^{a_h} \chi(q') \frac{a_h-q'}{a_h-a_l} dq' \\
&= 1 - \chi(a_h) \int_{a_l}^{a_h} \frac{2[a_h-a_l]^2 (a_h-q')}{\left[(a_h-q')^2+(q'-a_l)^2\right]^2} dq' \\
&= 1 - \chi(a_h) \left[-\arctan\frac{2x-a_l-a_h}{a_l-a_h} - \frac{(a_l-a_h)(x-a_l)}{(a_h-x)^2+(x-a_l)^2}\right] |_{a_l}^{a_h} \\
&= 1 - \chi(a_h)(2\arctan(1)+1) \\
&= 1 - \frac{1}{2+\frac{\pi}{2}}\left(\frac{\pi}{2}+1\right) \\
&= \frac{1}{2+\frac{\pi}{2}}.
\end{aligned}$$

Now we would like to derive closed forms for tuition functions:

$$
\begin{aligned}
t^{l}(q) &= \int_{a_l}^{q} \frac{y^{l}(q')}{\kappa+q} dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{1}{2}\Delta_y-\Delta_y\int_{q'}^{a_h}\chi(x)\frac{a_h-x}{a_h-a_l}dx\right)dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{1}{2}\Delta_y-\frac{\Delta_y}{2+\frac{\pi}{2}}\int_{q'}^{a_h}\frac{2\left[a_h-a_l\right]^2(a_h-x)}{\left[(a_h-x)^2+(x-a_l)^2\right]^2}dx\right)dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{1}{2}\Delta_y-\frac{\Delta_y}{2+\frac{\pi}{2}}\left(\begin{array}{c}-\arctan\dfrac{2x-a_l-a_h}{a_l-a_h}\\ -\frac{(a_l-a_h)(x-l)}{(a_h-x)^2+(x-a_l)^2}\end{array}\right)\Big|_{q'}^{a_h}\right)dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{1}{2}\Delta_y+\frac{\Delta_y}{2+\frac{\pi}{2}}\left(\begin{array}{c}\arctan\dfrac{2a_h-a_l-a_h}{a_l-a_h}+\frac{(a_l-a_h)(a_h-a_l)}{(a_h-a_h)^2+(a_h-a_l)^2}\\ -\arctan\dfrac{2q'-a_l-a_h}{a_l-a_h}-\frac{(a_l-a_h)(q'-l)}{(a_h-x)^2+(x-a_l)^2}\end{array}\right)\right)dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{1}{2}\Delta_y+\frac{\Delta_y}{2+\frac{\pi}{2}}\left(\begin{array}{c}-\frac{\pi}{4}-1-\arctan\dfrac{2q'-a_l-a_h}{a_l-a_h}\\ -\frac{(a_l-a_h)(q'-l)}{(a_h-q')^2+(q'-a_l)^2}\end{array}\right)\right)dq' \\
&= \frac{1}{\kappa+q}\int_{a_l}^{q}\left(\mu_y+\frac{\Delta_y}{2+\frac{\pi}{2}}\left(-\arctan\frac{2q'-a_l-a_h}{a_l-a_h}-\frac{(a_l-a_h)(q'-l)}{(a_h-q')^2+(q'-a_l)^2}\right)\right)dq' \\
&= \frac{1}{\kappa+q}\left[\int_{a_l}^{q} a\,dq'-\frac{\Delta_y}{2+\frac{\pi}{2}}\int_{a_l}^{q}\arctan\frac{2x-a_l-a_h}{a_l-a_h}+\frac{(a_l-a_h)(x-l)}{(a_h-x)^2+(x-a_l)^2}dx\right] \\
&= \frac{1}{\kappa+q}\left[\mu_y(q-a_l)-\frac{\Delta_y}{2+\frac{\pi}{2}}(x-a_l)\arctan\frac{2x-a_l-a_h}{a_l-a_h}\Big|_{a_l}^{q}\right] \\
&= \frac{(q-a_l)}{\kappa+q}\left[\mu_y-\frac{\Delta_y}{2+\frac{\pi}{2}}\arctan\frac{2q-a_l-a_h}{a_l-a_h}\right].
\end{aligned}
$$

Following similar steps:

$$
t^{h}(q) = -\frac{(a_h-q)}{\kappa+q}\left[\mu_y-\frac{\Delta_y}{2+\frac{\pi}{2}}\arctan\frac{2q-a_l-a_h}{a_l-a_h}\right].
$$

Rearranging and plugging in the expressions for $\eta(q)$, we obtain the theorem.

7.2 Computation Appendix

7.2.1 Algorithm for Baseline Calibration

This section explains the computational algorithm used to solve the quantitative model with two ability types. The key equilibrium object to solve for is the college distribution function $\chi(.)$ defined over a discrete grid on college quality q.

1. Construct a grid on college quality q.

2. Make an initial guess of the share of high-ability students not entering college: $\eta(0)$. By definition, the fraction of low-ability students not entering college is $1-\eta(0)$.

3. Solve for $\chi(0)$ from the zero profit condition of colleges of quality $q(1)$.

   (a) Starting with a conjecture for $\chi(0)$, compute the income of the "marginal" household attending colleges of quality $q(1)$:

   $$\begin{aligned} y\left(i^h(1)\right) &= y(\eta(0)\chi(0)) \\ y\left(i^l(1)\right) &= y((1-\eta(0))\chi(0)). \end{aligned}$$

   Next, pin down the college tuition $\left(t^h(1), t^l(1)\right)$ of the $q(1)$ college by the marginal household's indifference condition:

   $$\begin{aligned} \log\left(y\left(i^h(1)\right)\right)+\varphi\log(\kappa) &= \log\left(y\left(i^h(1)\right)-t^h(1)\right)+\varphi\log(\kappa+q(1)) \\ \log\left(y\left(i^l(1)\right)\right)+\varphi\log(\kappa) &= \log\left(y\left(i^l(1)\right)-t^l(1)\right)+\varphi\log(\kappa+q(1)). \end{aligned}$$

   Given the prevailing market tuition $t^h(1)$ and $t^l(1)$, solve the $q(1)$ college optimization problem (procedure outlined below) and obtain its profit $\pi(1)$ as well as its optimal decision rules $\eta(1), e(1)$.

   (b) Use the mapping described in part (a) to solve for $\chi(0)$ such that $\pi(1)=0$.

      i. Check the value of $\pi(1)$ at boundaries $\left[\chi_{lb}(0)=0; \chi_{ub}(0)=\min\left\{\frac{1}{\eta(0)}, \frac{1}{1-\eta(0)}\right\}\right]$. The upper bound arises because the total mass of high(low) ability is 1. Note that the profit $\pi(1)$ should be increasing in $\chi(0)$, as the market tuition rates $t^h(1)$ and $t^l(1)$ are both increasing in $\chi(0)$.

         A. If $\pi(1)>0$ at $\chi_{lb}(0)=0$, or $\pi(1)<0$ at $\chi_{ub}(0)=\min\left\{\frac{1}{\eta(0)}, \frac{1}{1-\eta(0)}\right\}$, zero profits cannot be obtained at grid $q(1)$. Thus we delete $q(1)$ and set $q(1)=q(2)$ and go back to step 3, else go to step ii.

      ii. As $\pi(1)<0$ at $\chi_{lb}(0)=0$ and $\pi(1)>0$ at $\chi_{ub}(0)$, one can solve for $\chi(0)$ from $\pi(1)=0$ using a simple one-dimensional nonlinear solver.

4. Having solved for $\{\chi(i)\}_{i=1}^{n-1}$, along with $\{\eta(i), e(i)\}_{i=1}^{n}$ we now solve for $\chi(n)$ from $\pi(n+1)=0$.

(a) Starting from a conjecture for $\chi(n)$, compute the income of the marginal household attending colleges of quality $q(n+1)$:

$$\begin{aligned} y\left(i^h(n+1)\right) &= y\left(\sum_{i=0}^{n} \eta(i)\chi(i)\right) \\ y\left(i^l(n+1)\right) &= y\left(\sum_{i=0}^{n} (1-\eta(i))\chi(i)\right). \end{aligned}$$

Next, pin down the college tuition $\left(t^h(n+1), t^l(n+1)\right)$ of the $q(n+1)$ college from the following household first-order conditions:

$$\begin{aligned} t^h(n+1) &= \left[1-\left(\frac{\kappa+q(n)}{\kappa+q(n+1)}\right)^{\varphi}\right] y\left(i^h(n+1)\right) + \left(\frac{\kappa+q(n)}{\kappa+q(n+1)}\right)^{\varphi} t^h(n) \\ t^l(n+1) &= \left[1-\left(\frac{\kappa+q(n)}{\kappa+q(n+1)}\right)^{\varphi}\right] y\left(i^l(n+1)\right) + \left(\frac{\kappa+q(n)}{\kappa+q(n+1)}\right)^{\varphi} t^l(n). \end{aligned}$$

Given the prevailing market tuition $t^h(n+1)$ and $t^l(n+1)$, solve the $q(n+1)$ college optimization problem (procedure outlined below) and obtain its profit $\pi(n+1)$ as well as the optimal decision rules $\eta(n+1), e(n+1)$.

(b) Given the mapping described in part (a), solve for $\chi(n)$ such that $\pi(n+1)=0$.

i. Check the value of $\pi(n)$ at boundaries $\left[\chi_{lb}(n)=0; \chi_{ub}(n)=\min\left(\frac{1-i^h(n)}{\eta(n)}, \frac{1-i^l(n)}{(1-\eta(n))}\right)\right]$.

A. If $\pi(n+1)>0$ at $\chi_{lb}(n)$, this implies that $q(n+1)$ college would always make strictly positive profits and keep growing, squeezing $q(n)$ college out of the market. Thus, we can delete the grid point $q(n)$. Set $q(n)=q(n+1)$ and go back to step 3 with the new grid on q.

B. If $\pi(n+1)<0$ at $\chi_{ub}(n)$, this implies that $q(n+1)$ college always makes negative profits and thus is driven out of the market. Thus, we delete the grid point $q(n+1)$. Set $q(n+1)=q(n+2)$ and go back to step 3 with the new grid on q.

C. If $\pi(n+1)<0$ at $\chi_{lb}(n)$ and $\pi(n+1)>0$ at $\chi_{ub}(n)$, then we can solve for $\chi(n)$ such that $\pi(n+1)=0$.

5. Having solved for $\{\chi(i)\}_{i=1}^{N-1}$, along with $\{\eta(i), e(i)\}_{i=1}^{N}$, we still have $\chi(N)$ undetermined. We pin it down using the consistency requirement for high ability spots at $q(N)$ colleges:

$$\chi(N)\eta(N) = 1 - i^h(N).$$

Lastly, check the consistency requirement for low ability spots at $q(N)$ college:

$$1 - i^l(N) - \chi(N)(1-\eta(N)) = 0.$$

If this requirement is satisfied (to desired numerical accuracy), stop. If not, go back to step 2 and adjust $\eta(0)$.

Subalgorithm for solving the individual college's problem: We now describe the algorithm to solve colleges' optimization problem given that they deliver quality q and can charge tuition t^h and t^l for high- and low-ability students, respectively, taking into account the tuition and quality bounds as well as the subsidies.

1. We first solve for college profit if the college chooses to become private (type 2). To compute profit, we need to consider two cases:

   (a) If $t^h \geq t^l$, then the college admits only high ability students. Thus,

   $$\begin{aligned} a_2^* &= a_h \\ \eta_2^* &= \frac{a - a_l}{a_h - a_l} \\ e_2^* &= q^{\frac{1}{1-\theta}} \left(a_2^*\right)^{-\frac{\theta}{1-\theta}}, \end{aligned}$$

   and the profit in this case is given by

   $$\pi_2 = \eta_2^* t^h + \left(1 - \eta_2^*\right) t^l - e_2^* - \phi + s_2 - \omega.$$

   (b) If $t^h < t^l$, then its optimal decision rules are given by

   $$\begin{aligned} a_2^* &= q \left(\frac{(1-\theta)\left(t^l - t^h\right)}{\theta\left(a_h - a_l\right)} \right)^{\theta - 1} \\ \eta_2^* &= \frac{a - a_l}{a_h - a_l} \\ e_2^* &= q^{\frac{1}{1-\theta}} \left(a_2^*\right)^{-\frac{\theta}{1-\theta}}, \end{aligned}$$

   and profit in this case is given by

   $$\pi_2 = \eta_2^* t^h + \left(1 - \eta_2^*\right) t^l - e_2^* - \phi + s_2 - \omega.$$

2. We next solve for college profit if the college chooses to become public. We need to consider the following six cases:

   (a) If $t^h \leq t^l \leq \bar{T}$. Then the public college is unconstrained by the tuition cap. Thus,

   $$\begin{aligned} a_1^* &= q \left(\frac{(1-\theta)\left(t^l - t^h\right)}{\theta\left(a_h - a_l\right)} \right)^{\theta - 1} \\ \eta_1^* &= \frac{a - a_l}{a_h - a_l} \\ e_1^* &= q^{\frac{1}{1-\theta}} \left(a_1^*\right)^{-\frac{\theta}{1-\theta}}, \end{aligned}$$

   and profit is given by

   $$\pi_1 = \eta_1^* t^h + \left(1 - \eta_1^*\right) t^l - e_1^* - \phi + s_1 - \omega.$$

(b) If $t^h \leq \bar{T} \leq t^l$, then it is possible that the college is constrained by the tuition cap.

i. Solve for the decision rule ignoring the tuition cap:

$$\begin{aligned} a_1^* &= q\left(\frac{(1-\theta)\left(t^l-t^h\right)}{\theta\left(a_h-a_l\right)}\right)^{\theta-1} \\ \eta_1^* &= \frac{a-a_l}{a_h-a_l} \\ e_1^* &= q^{\frac{1}{1-\theta}}\left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

ii. Check if the tuition cap is violated: $\eta_1^* t^h+(1-\eta_1^*)\,t^l<\bar{T}$. If not, proceed to case c. If so, the decision rule should respect the tuition cap and is given by

$$\begin{aligned} \eta_1^* &= \frac{t^l-\bar{T}}{t^l-t^h} \\ a_1^* &= \eta_1^* a_h+\left(1-\eta_1^*\right) a_l \\ e_1^* &= q^{\frac{1}{1-\theta}}\left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

Profit is given by

$$\pi_1=\eta_1^* t^h+\left(1-\eta_1^*\right) t^l-e_1^*-\phi+s_1-\omega.$$

(c) If $\bar{T} \leq t^h \leq t^l$, then the college can charge at most $\bar{T}$ to each student. Thus, it will admit only the high-ability students:

$$\begin{aligned} \eta_1^* &= 1 \\ a_1^* &= a_h \\ e_1^* &= q^{\frac{1}{1-\theta}}\left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

Profit is given by

$$\pi_1=\bar{T}-e_1^*-\phi+s_1-\omega.$$

(d) If $t^l \leq t^h \leq \bar{T}$, then admit only the high ability students

$$\begin{aligned} \eta_1^* &= 1 \\ a_1^* &= a_h \\ e_1^* &= q^{\frac{1}{1-\theta}}\left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

Profit is given by

$$\pi_1=\eta_1^* t^h+\left(1-\eta_1^*\right) t^l-e_1^*-\phi+s_1-\omega.$$

(e) If $t^l \leq \bar{T} \leq t^h$, then the college is constrained by the tuition cap and will charge tuition equal to $\bar{T}$:

$$\begin{aligned} \eta_1^* &= 1 \\ a_1^* &= a_h \\ e_1^* &= q^{\frac{1}{1-\theta}}\left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

Profit is given by

$$\pi_1 = \bar{T} - e_1^* - \phi + s_1 - \omega.$$

(f) If $\bar{T} \leq t^l \leq t^h$, Then the college can charge at most $\bar{T}$ to each student. Thus, it will admit only the high-ability students:

$$\begin{aligned} \eta_1^* &= 1 \\ a_1^* &= a_h \\ e_1^* &= q^{\frac{1}{1-\theta}} \left(a_1^*\right)^{-\frac{\theta}{1-\theta}}. \end{aligned}$$

Profit is given by

$$\pi_1 = \bar{T} - e_1^* - \phi + s_1 - \omega.$$

3. Compare profits. If the quality bound is violated $q < Q$ or private college is more profitable, $\pi_2 > \pi_1$, then set $\pi = \pi_2$. Otherwise, set $\pi = \pi_1$ This concludes the sub-algorithm for solving individual college's problem.

7.2.2 More Ability Types: Computation

This appendix has two parts. In the first part, a computational algorithm is sketched and can be used to solve a model with more than two ability types.[36] In the second part, we compare the quantitative results in our baseline calibration with 2 ability types to a 10-ability-type case, which we solve using the algorithm proposed in the first part. We find that varying the number of grid points has negligible effects on the equilibrium quality distribution and on key moments of the enrollment and tuition distributions.

Here, we abstract from the distinction between public and private colleges by assuming no quality bounds or tuition thresholds. This eliminates the "holes" in the quality distribution that arise in our baseline quantitative model and ensures that zero entry in the market can arise only at the bottom of the college distribution (as in the baseline model, zero entry here reflects the presence of fixed costs). Thus, the strategy is to solve for a college distribution over a quality grid where we know with certainty that colleges are active, and to then check for profitable entry at the bottom.

1. Set up a grid of college quality $(q_1, q_2, ..., q_N)$ where we know with certainty that colleges enter.

2. Make an initial guess of the vector of corresponding discount rates $D_0 = (d_1, d_2, ..., d_N)$.

3. Given the discount rates, use the college first-order conditions and zero profit conditions to compute the set of implied baseline tuition $(b_1, b_2, ..., b_N)$. Thus, we obtain a full set of tuition schedules.

4. Given the tuition schedules, use the household's indifference condition to pin down a set of income thresholds for each ability type $(y_1^a, y_2^a, ..., y_N^2)$, where y_i^a is the income of a household indifferent between quality $i-1$ and quality i colleges.

[36]More details are available upon request.

Figure 11: College Distribution

compare2_10.pdf

5. Given the income thresholds, compute the supply of average ability to each college $(a_1^s, a_2^s, ..., a_N^s)$.

6. Given demand for ability by each college i (pinned down given d_i from the college first-order conditions) $(a_1^d, a_2^d, ..., a_N^d)$ check market clearing: $a_n^d - a_n^s = 0\ \forall n$.

7. If markets do not clear, go back to step 2 and adjust the discounts d_n.

8. Check for profitable entry at the bottom. For instance, suppose a college of quality $q_0 < q_1$ enters. To charge the maximum tuition, it has to appeal to the marginal household with income y_1^a. Thus, we can use the household's indifference condition to pin down tuition t_0^a. We then solve the college problem and check its profit. If profit is positive, go back to step 1 and add q_0 to the grid of college quality. Otherwise, we stop.

7.2.3 More Ability Types: Quantitative Comparison

We now solve the college model with 10 points in the grid on ability and compare the results to our benchmark two-ability-types calibration. We use the same parameterization as in our baseline. We discretize the 10 grid points such that 1) the 10-grid-point model has the same variance of ability as the baseline, and 2) the conditional mean of income distribution varies linearly with ability with the same slope as the baseline.

Table 8 reports various first and second moments for the baseline and the 10-ability-type model. Note that because there are no public colleges here, the statistics are slightly different from those in Table 3. We also plot the college distribution function χ across the two cases. All the results suggest that moving to a finer ability grid has a negligible impact on college enrollment and pricing patterns.

Table 8: Comparison

Statistics	Baseline	10 Ability Types
First Moments		
Enrollment	36.9%	36.9%
Average tuition	0.0479	0.0479
Average college quality/κ	4.721	4.727
Average income	2.015	2.015
Average ability	0.821	0.821
Average discount	0.0305	0.0304
Second Moments		
Standard Deviation/ Mean		
Tuition	0.875	0.865
Quality	0.028	0.028
Family income	1.399	1.389
Ability	0.078	0.081
Correlation		
Tuition vs. quality	0.922	0.924
Tuition vs. ability	0.455	0.484
Tuition vs. income	0.999	0.999
Ability vs. income	0.465	0.463

7.3 Data Appendix

In this data appendix, we describe in detail how we construct the empirical moments presented in the main text.

Tables 1 and 2

College attendance: Share of population 25–34 with four-years college or more (2016) from US Census Bureau, Education and Social Stratification Branch, Table A-1.

Share of students in public and private colleges: College Board (2016a), Figure 20.

Sticker tuition: College Board (2016a), Table 7.

Net tuition: College Board (2016a), Table 7.

Federal and state grant aid: Federal and state grant aid in 2014 was 15.0 percent of total grant aid at private non-profits (College Board (2016b), Figure 15). Thus, the federal and state grant aid at those private nonprofit schools is 15.0 percent $*(33,480-14,190) = 2893.5$. Federal and state grant aid was 54.4 percent of total grant aid at public schools (College Board (2016b), Figure 15). Thus, the grant aid at public schools is 54.4 percent $*(9,650-3,770) = \$3,198.7$.

Institutional Aid: Institutional aid is equal to total aid (sticker tuition less net tuition) less federal and state grant aid. For private nonprofit schools, it is $33,480-14,190-2,893 = \$16,397$. For public schools, it is $9,650-3,770-3,199 = \$2,681$

Forgone earnings: From the CPS (Series ID: LEU0252886300) we take median usual weekly earnings for full-time wage and salary workers aged 16-24. In current dollars, this was \$259 in 1990 and \$501 in 2016. Adjusted by the CPI and assuming 40 weeks of college per year, forgone earnings from attending college is \$20,040 in 2016 dollars in both 2016 and 1990.

Room and board: College Board (2016a), Table 7.

Instructional spending and student services: These data are from NCES (National Center for Education Statistics) Digest of Education Statistics (2017 tables). For public colleges, see Table 334.10, columns C and H. For private colleges, see Table 334.30, columns C and G.

Table 3

In the first moments section, graduation, net tuition, and sticker tuition are all taken from Tables 1 and 2. The rest of the statistics including the second moments are computed from College Scorecard microdata merged with the Mobility Report Cards data set (Chetty et al. 2017), which has higher-quality household income data. Specifically, we download the most recent data from the College Scorecard (https://collegescorecard.ed.gov/data/) and merge it with household income data from the Mobility Report Cards online data Table 2.[37] Next, we describe how we construct each variable.

Sticker tuition: In-state tuition and fees (variable name tuitionfee_in).

Net tuition: Average net price paid (NPT4_pub for public colleges and NPT4_priv for private colleges). Note this measure includes the full cost of attendance (including living expenses). We construct a measure of living expenses by subtracting tuition and fees (tuitionfee_in) from the full cost of attendance (costt4_a). Net tuition is then obtained by subtracting the living expenses from the average net price paid.

[37] Note that to link online data Table 2 to College Scorecard data, we need to use online data Table 11.

Household income: From Mobility Report Cards online data Table 2 (variable name par_mean). Mean income is $87,335 (Chetty et al. 2017, Table 2).

Fraction of high ability: We collect data on national averages of the SAT score (whenever the SAT score is not available, we substitute the ACT score.) We then assume that the score is normally distributed at each college and use the college-specific 25th percentile and 75th percentile SAT score (satmt25, satvr25, satwr25, satmt75, satvr75, satwr75) to back out the mean and variance of the distribution at each college. Then we compute the fraction of high-ability students at each college as the fraction with a score higher than the national average.

Table 5

We target exactly the same empirical moments for the 1990 calibration as in the 2016 calibration. We infer the price of expenditure p from the growth of faculty salaries.

College attendance: 24.2 percent. Share of population 25–34 with four years' college or more taken from US Census Bureau, Education and Social Stratification Branch, Table A-1.

Net tuition: $11,750 for private nonprofit colleges and $2,000 for public colleges(College Board (2016a), Table 7).

Faculty salaries: $75,024 in 1990 and $82,101 in 2016 (NCES Table 316.10).

Share of students in public colleges: 70.69 percent (NCES Digest of Education Statistics, Table 303.70)

Variable expenditure: We want to measure variable expenditure (defined as instructional expenses+student services) in 1990 in the same way as in 2016. The difficulty is that the NCES changed its reporting standards twice during late 1990s and early 2000s, making numbers not directly comparable across 1990 and 2016. We instead compute growth rates in each subperiod during which reporting standards remained consistent, and use these growth rates, together with the 2014-2015 number, to infer 1990 expenditure.

For public colleges, the cumulative growth of variable expenditure was 17.9 percent from 1990 to 2001.[38] The growth rate was 16 percent from 2003 to 2014 (College Board, 2016a). Given that the variable expenditure for public colleges was $11,881 in 2014-2015, we infer that consistently defined variable expenditure in 1990 was $8,680.

For private nonprofit colleges, the growth rate was 11.3 percent between 1990 and 1996.[39] The growth rate was 10.44 percent between 1996 and 1999.[40]. The growth rate was 7.44 percent between 1999 and 2003.[41] The growth rate between 2003 and 2014 was 13.5 percent (College Board, 2016a). Given these growth rates and that the variable expenditure for private colleges was $22,120 in 2014-2015, the variable expenditure for private nonprofit colleges in 1990 is $14,757.

[38] See NCES Digest of Education Statistics, 2004, Tables 347 and 348. In this period, NCES divides public colleges into two groups: public universities and public four-year colleges. For public universities, it was $8,158 in 1990 and $9,604 in 2001, while for public four-year colleges it was $7,035 in 1990 and $8,299 in 2001. The growth rate was the same for these two types of public colleges.

[39] NCES Digest of Education Statistics, 1999, Tables 352 and 353, Table 176: share of private universities is 0.2796 in 1990. The average expenditure in 1990 is 0.2796*13,274+(1-0.2796)*6,381=$8,308.3. The share of private universities in 1996 is 0.2614. In 1996 average expenditure is 0.2614*15,298+(1-0.2614)*7,105=$9,246.7.

[40] NCES Digest of Education Statistics, 2006, Table 352.

[41] See NCES Digest of Education Statistics, 2016, Table 334.30.