The Consumption Value of Postsecondary Education[*]

Brian Jacob
University of Michigan and NBER

Brian McCall
University of Michigan

Kevin Stange
University of Michigan

November 2, 2011

Abstract

Education provides both investment and consumption benefits; the former being realized after schooling is completed but the latter accruing only while schooling is actually taking place. In this paper, we quantify the importance of consumption value considerations to schooling decisions in the context of higher education and examine the implications for colleges' strategies for attracting students. To do so, we estimate a discrete choice model of college demand using micro data from the high school classes of 1992 and 2004, matched to extensive information on all four-year colleges in the U.S. We find that most students do appear to value college attributes which we categorize as "consumption," including college spending on student activities, sports, and dormitories. In fact, students appear to be more willing to pay for these non-academic aspects of colleges than typical academic aspects, such as spending on instruction. Estimates suggest that this taste for consumption amenities is broad-based among many student groups, whereas taste for academic quality is confined only to the high achieving. Consequently, policies that reallocate financial resources away from these non-academic aspects to instruction would not enable most schools to attract more or better students, as some policy-makers suggest. However, since student preferences for college attributes are very heterogeneous different colleges face very different incentives for changing their characteristics depending on their current student body and those they are trying to attract.

Our empirical approach makes a number of improvements on existing literature, including accounting for unobserved choice set variability created by selective admissions, controlling for fixed unobserved differences between schools and price discounting, and permitting greater preference heterogeneity.

[*] We would like to thank Jon Bartik, Michael Gelman, Jonathan Hershaff, Max Kapustin, Geoff Perrin, Nathan Schwartz, Elias Walsh, and Tamara Wilder for their research assistance. We would also like to recognize the valuable comments and suggestions provided by various seminar participants. Jacob can be contacted at bajacob@umich.edu; McCall can be contacted at bpmccall@umich.edu; Stange can be contacted at kstange@umich.edu. All errors are our own.

"Boys and girls and their parents too often choose an educational institution for strange reasons: because it has lots of outdoor life; a good football team; a lovely campus; because the president or the dean or some professor is such a nice man" (Tunis 1939, p. 7).

1. Introduction

Economists' treatment of education traditionally employs the human capital framework developed by Becker (1964). In this framework, education is viewed primarily as an investment wherein individuals forgo current labor market earnings and incur direct costs in return for higher future wages. The original theoretical work by Becker (1964), Ben Porath (1967) and others spurred a tremendous amount of empirical work, which has generally supported the implications of the human capital model (Freeman 1986).[1] At the same time, the human capital framework does not rule out that education may also provide immediate consumption. Indeed, many economists have discussed the consumption value of education. For example, Schultz (1963) identifies current consumption as one of three benefits of education, along with investment and future consumption. More recently, Oreopoulos and Salvanes (2011) highlight consumption considerations in their recent review of the non-pecuniary returns to education. For the most part, however, consumption aspects of education have received relatively little attention in the literature.

Several trends suggest that consumption may be becoming an increasingly important part of the choice of whether, where, and how to attend college. Babcock and Marks (forthcoming) document that academic time investment among full-time college students fell from 40 hours per in 1961 to 27 hours per week in 2003 and that this change cannot be explained by compositional changes or changes in work and major choices. At the same time, some observers have argued that increased market pressure has caused colleges to cater to students' desires for leisure (Kirp 2005). Recent analysis by the Delta Cost Project (2010) found that colleges' spending on student services has outpaced that on instruction for the past decade for all postsecondary sectors. Bound et al. (2007) document that college completion rates have declined, primarily among men entering two-year or less selective four-year institutions. Low completion despite high returns to earning a degree is consistent with schooling serving primarily as consumption rather than

[1] The model implies that the demand for education increases with the years of potential employment, the rate of return of education, and the efficiency with which the individual can translate time and money into human capital. Conversely, the demand for education will decrease with an individual's discount rate.

investment. Scott-Clayton (2007) shows that college students are working substantially more today than in the past, and that the increased labor supply cannot be easily explained by compositional changes in students or schools, or credit constraints.[2] These trends have put what colleges do and what students actually learn while there in the spotlight. For instance, the tension universities' participation in big-time commercial sports creates for their traditional academic values was the subject of Clotfelter (2011) and the extent of learning occurring during college was examined by Arum and Roksa (2011).

Educational choices may thus mirror a broader shift towards present-day consumption and away from investment. Recent macroeconomic literature has documented that the share of output that is consumed has increased dramatically since the early 1980s, reducing the personal savings rate.[3] Just as less investment creates concern about the future productivity of America's physical capital, a greater share of educational resources devoted towards (non-productive) immediate consumption creates concerns about the future productivity of America's workforce.

In this paper, we attempt to more carefully quantify the importance of direct consumption aspects of postsecondary education, how this varies across students, and what implications this has for colleges. As we describe in more detail below, by consumption we mean the direct and immediate utility that one derives while enrolled from attending a particular type of schooling or a particular institution.[4] In contrast to previous work on the subject, our approach is to infer demand for consumption aspects of education from students' college choices. Just as colleges are able to attract students by increasing their academic reputation, students' responses to colleges' social aspects signal a willingness to pay for consumption aspects. Our main approach is to estimate a discrete choice model of college demand using micro data from the high school classes of 1992 and 2004, matched to extensive information about the universe of nearly all four-year colleges in the U.S. This approach is in the spirit of the standard differentiated product

[2] To the extent that college-related leisure and current income are complementary, this result is consistent with an increasing importance of consumption.

[3] Possible explanations include a wealth effect from capital gains, changes in the discount rate and preferences, an increase in consumption propensities of older Americans, relaxed liquidity constraints, and government's redistribution of resources toward older generations with high consumption propensities from younger ones. See Parker (1999); Gokhale, Kotlikoff, Sabelhaus (1996); Gale and Sabelhaus (1999), Poterba (2000), Juster, Lupton, Smith, and Stafford (2005), Greenspan and Kennedy, (2007).

[4] Some related literature describes the benefits that education confers on subsequent household production as a "non-monetary" or "consumption aspect" of education in the sense that it increases the efficiency of future consumption (see Michaels 1973 for a discussion of the education and household production). These benefits of education would not count as consumption value in our framework as they accrue post-schooling.

demand models recently used to study product demand (e.g. Berry, Levinsohn, Pakes, 1995), residential choice (e.g. Bayer, Ferreira, McMillan, 2007), and school choice (e.g. Hastings, Kane, and Staiger, 2009), among others. In this approach, preference parameters are inferred from observed college choices, where each college is a bundle of characteristics.

We find that students do appear to value several college attributes which we categorize as "consumption" because their benefits arguably accrue only while actually enrolled. For instance, college spending on student activities, sports, and dormitories are significant predictors of college choice, and more influential than instructional spending and academic support. Specifically, we estimate that students would be willing to pay 7 percent more to attend a school that spends 100 percent more on student and auxiliary services (dorms, sports, and food service) but is unwilling to spend more to attend a college that spends more on instruction (in fact the point estimate is negative). However, there is significant heterogeneity of preferences across students, with higher achieving students having a greater willingness-to-pay for academic quality than their less academically-oriented peers. The existence of significant preference heterogeneity has important implications for how colleges can attract students. Policies that reallocate financial resources away from these non-academic aspects to instruction would not enable most schools to attract more or better students, as some policy-makers suggest. However, since student preferences for college attributes are very heterogeneous different colleges face very different incentives for changing their characteristics depending on their current student body and those they are trying to attract.

While it is not obvious ex-ante that these spending categories are good proxies for "consumption" vs. "academic" amenities of schools, students' self-reported preferences are consistent with this interpretation. High school students who list "social environment" as an important factor in their college decision are more likely to attend colleges that spend more on what we categorize as consumption whereas students who list "academic reputation" as a top priority are more likely to attend schools that spend relatively more on instruction. This pattern is robust to controls for admission difficulty and unobserved college characteristics.

Our analysis makes several contributions to the existing literature. First, we explicitly examine the role of non-academic aspects of the college experience in student choice. Second, our estimation strategy addresses several important shortcomings in much prior work on college choice. Specifically, we (1) explicitly account for unobserved choice set variability created by selective admissions; (2) use variation over time to identify the importance of institutional characteristics that are fixed within a specific college-cohort (e.g., per pupil spending on instruction) while controlling for fixed unobserved differences between schools and also address price discounting; and (3) incorporate substantial preference heterogeneity that relaxes the IIA assumption to allow more flexible substitution patterns across institutions. Our analysis also examines a more recent cohort of students than previous work.

The remainder of the paper proceeds as follows. The next section reviews the prior literature on the consumption value of education and on college choice. Section 3 presents a simple model of college choice in order to make explicit the parameters we are interested in estimating. Section 4 introduces our empirical strategy and elaborates on the identification challenges. Our data sources are discussed in Section 5. Our main results are presented in Section 6. Section 7 discusses the implications of student demand for colleges. Section 8 concludes.

2. Prior literature

In this section, we review the prior attempts to estimate the consumption value of postsecondary education as well as the extensive college choice literature.

2.1 Previous Literature on the Consumption Value of Education

Prior studies have examined the consumption value of education in two different ways: (1) comparing the total amount of education one obtains to the income maximizing one; and (2) examining the type of degree (or major) one pursues.

The first line of inquiry has focused on the total amount of education attained. These papers seek to estimate the financially optimal amount of schooling for individuals, and then compare it to the observed level of schooling attained. If individuals consume more schooling than is optimal from a purely financial perspective, then one would conclude that schooling itself must contribute directly to utility. In one of the first papers to take this approach, Lazear (1977) develops a model to test whether the observed relationship between income and education reflects a causal impact of education. To do so, he develops a model of education that incorporates both investment and consumption goods. He then estimates the parameters of this model using data from the National Longitudinal Survey, which follows 14-24 year olds starting in 1966. He finds that individuals obtain less than their wealth-maximizing level of education, suggesting that education actually has a negative consumption value – i.e., it is a bad.[5]

Kodde and Ritzen (1984) develop a similar model of educational attainment that allows schooling to have a direct, positive impact on utility. They start with the observation that many studies show positive income effects of education, which is ruled out in the case of the standard model and perfect capital markets. They also note that studies find different enrollment effect of forgone earnings and direct tuition costs, which have been explained in terms of capital constraints or measurement error. However, they point out that these differential enrollment effects are consistent with the model they develop. Specifically, an increase in wage rate will reduce enrollment less than increase in the direct cost of schooling because higher wages imply greater wealth and people will choose to consume some of this greater wealth by buying more education.

Oosterbeek and van Ophem (2000) write down a structural model of the determinants of schooling attainment that allows utility to be a function of future earnings as well as schooling itself. The model includes two simultaneously determined outcomes: years of schooling and log wage rate. They model the preference for schooling as a function of a student's grades in the final year of compulsory school and the parent's interaction with the child's school. They model discount rates as a function of father's education and occupation. They include student IQ and a quadratic in years of experience and a quadratic in schooling in the wage equation directly. In addition to functional form, their model is identified by assuming that the father's education, the student's grades and the parent's interaction with the child's school only influence schooling and do not enter the wage equation directly. They find that the school preference parameter is non-zero and depends positively on student grades and family social status, suggesting that consumption is a significant determinant of educational attainment.

A related approach is exemplified by Heckman et al (1999) and Carniero et al (2000). They attempt to quantify the psychic benefits and costs of attending college. Using data on male earnings in the United States, Heckman et al (1999) find that individuals in the second-highest ability quartile enjoy large nonpecuniary benefits from attending college; individuals in the other quartiles suffer non-pecuniary costs. Carneiro et al. (2003) estimate that, when ignoring psychic

[5] The paper is not really directly focused on consumption aspects, at least in terms of how we are thinking of them. The approach describes income effects in education as consumption value.

gains, forty percent of college attendees would regret it. Once they account for psychic benefits and costs of attending college, only 8 percent of college graduates regret attending college. The authors conclude, therefore, that much of the gain from college is nonpecuniary.

A second strand of research focuses on the type of degree (or major) that students choose. The general intuition in these papers is that an individual's decision to obtain a degree with a significantly lower long-term financial return than the individual could have obtained in another field (given the individual's observed ability) provides evidence that schooling (or at least certain degrees) have direct utility value. For example, Alstadsæter (2009) estimates that individuals who attended Teacher's College in Norway during the 1960s gave up substantial future wages to do so. She calculates that their willingness-to-pay for the teaching degree (relative to a business degree) was roughly 35 percent of the present value of their potential lifetime income. However, this study cannot distinguish between an individual's preference for a particular type of work and a preference for a particular type of academic experience.

Arcidiacono (2004) develops a more comprehensive model of student choice of institution type and college major that allows for both direct and immediate utility effects of the type of schooling (i.e., the "costs" of studying a particular field in a particular institution) and direct (but future) utility effects of working in a particular occupation. He finds large differences in wage returns across college majors, even conditional on student ability. He concludes that preferences for different educational types are critical to decision-making. Importantly, preferences for studying a particular field in college are critical.

These approaches are not able to separate an individual's preference for a particular type of work from a preference for college itself. The choice to attend college implies a particular career path, which incorporates not only monetary rewards, but different working conditions and, indeed, a different "type" of work that may provide different direct utility to individuals. The same is true in the case of a college major. For example, the choice to major in engineering instead of education influences not only how the individual will spend their college years, but also the type of work one will do for the duration of their career.

2.2 Previous Literature on College Choice

Our approach deviates from this existing literature by using the college attributes demanded by students and supplied by institutions to identify preferences for consumption.[6] Empirical models of college choice have a long history, exemplified by the seminal work of Manski and Wise (1983). In general, discrete choice models of college enrollment have focused on estimating the importance of price, academic quality and distance. In perhaps the most thorough application of this approach, Long (2004) estimates a conditional logit model using data on high school graduates in 1972, 1982 and 1992. She finds that the role of college costs decreased over this period, and were not a significant factor in the decision to enroll, though it continues to be a significant factor in the decision of where to enroll. Distance also became less important while proxies for college academic quality such as instructional expenditures per student became more important over time. More recently, McDuff (2007) exploits cross-state variation in the cost and quality of public flagship universities and estimates that students' willingness to pay for academic quality is large. There are also a number of papers that use a reduced form approach to estimate the effect of academic quality or reputation (as measured by USWNP rankings) on number and quality of applicants and student yield. Typical is Monks and Ehrenberg (1999) who find that a ranking decline leads institutions to accept more of its applicants, have a lower matriculation rate among admitted students, enroll lower-ability students (as measured by average SAT scores), and decrease net tuition.

These models have not traditionally examined college consumption amenities. However, the recognition that college choice depends on a variety of factors beyond investment is not new. Writing in 1939, Tunis remarks that "Boys and girls and their parents too often choose an educational institution for strange reasons: because it has lots of outdoor life; a good football team; a lovely campus; because the president or the dean or some professor is such a nice man" (Tunis 1939, p. 7). Various studies since then have identified social considerations as an important factor in the college choice decision (Bowers and Pugh 1973, Keller & McKewon, 1984; Stewart, et al., 1987; Chapman & Jackson, 1987, Weiler 1996, Rosenbaum, Miller and Krei 1996).

While most of the research that focuses on social considerations is qualitative in nature, several studies have attempted to estimate the importance of such factors. For example, Weiler

[6] Our approach is somewhat related to the approach of Jacob and Lefgren (2007). They find that wealthy parents want teachers that both teach and increase student satisfaction. This latter aspect could be considered "consumption value" in our framework.

(1996) analyzes the matriculation decisions of a sample of high ability students who were admitted to a single selective research university. The researchers administered a survey to the top half of the admitted class as measured by academic aptitude. The survey asked students to rate the survey institution and up to two others to which s/he was admitted on a variety of non-monetary characteristics such as the quality of the social life and the academic reputation. Students were also asked about the financial aid and costs of attendance at each school. Finally, the survey asked students to rank his or her top choices. Using the information, the author estimates discrete choice models to assess the importance of various school attributes. Weiler finds that attendance costs and non-monetary institutional characteristics are both significant determinants of institutional choice. Among the non-monetary characteristics, those associated with non-academic items like housing and recreational options have about the same impact as academic concerns such as availability of majors or concentration on undergraduate education. [7]

Using a panel of NCAA Division 1 sports schools, Pope and Pope (2008, 2009) find that football and basketball success increases the quantity of applications colleges receive and the number of students sending SAT scores. Since the additional applications come from both high and low SAT scoring students, colleges are able to increase both the number and quality of incoming students following sports success.

To summarize, while there is ample evidence on the responsiveness of college decisions to academic and cost attributes of colleges, there is virtually no evidence on the importance of consumption considerations or how this importance has changed over time. This paper attempts to fill this gap.

3. Choice model and parameters of interest

In this section we describe a simple model of college demand in the spirit of the standard differentiated product demand models recently used to study product demand (e.g. Berry, Levinsohn, Pakes, 1995), residential choice (e.g. Bayer, Ferreira, McMillan, 2007), and school choice (e.g. Hastings, Kane, and Staiger, 2009). In this approach, preference parameters are recovered from observed college choices, where each college is a bundle of characteristics.

Individuals have J total colleges to choose from, each with a variety of different attributes. We partition college characteristics into those that are primarily oriented towards

[7] Chapman and Jackson (1987) explore similar factors. Lin, L. (1997) and Donnellan, J. (2002) are marketing studies that also are relevant. Also see Drewes and Michael (2006).

academic pursuits (i.e., investment) versus those that are more related to current consumption while in school. For instance, we think of colleges' instructional spending and the quality of peers as academic attributes, while intercollegiate sports spending and good weather are consumption amenities. In a later section, we more carefully describe the college characteristics used in the analysis.

Individuals receive indirect utility from attending college *j* that is separable in these two dimensions (denoted by A_j and C_j, respectively) and consumption of all other goods ($Y_i - T_{ij}$) where Y_i is income and T_{ij} is the price of college *j* to individual *i*. Individuals also care about the distance from their home to college *j*, D_{ij}, a proxy for the non-monetary commuting costs. Indirect utility is given by:

$$U_{ij} = \alpha_{1i}(Y_i - T_{ij}) + \alpha_{2i}A_j + \alpha_{3i}C_j + \alpha_{4i}D_{ij} + \varepsilon_{ij} \quad (1)$$

where ε_{ij} is an unobserved individual-specific taste preference for school *j*. Individuals compare the potential utility received from attending each college and choose to attend the college that maximizes their utility.

We are interested in estimating the coefficients a_{1i}, a_{2i}, a_{3i}, and a_{4i}, which correspond to the marginal utility individual *i* receives from each of the four college attributes. Since the absolute level of these coefficients does not matter, we focus instead on ratios between these coefficients as measures of the willingness to trade-off one characteristic for another. For instance, we interpret a_{2i}/a_{1i} as student *i*'s willingness to pay in dollars (WTP) for a one unit increase in academic quality. The ratio a_{3i}/a_{2i} is the rate at which student *i* could trade academic quality for consumption amenities and maintain a constant utility.

4. Empirical Strategy

Our objective is to estimate the parameters of (1) in order to calculate the willingness to pay for attributes that reflect direct consumption amenities separate from those for academic quality. To do so, we estimate a discrete choice model of college choice, taking the supply of college attributes as exogenous. Our approach builds on that of Long (2004), but extends her conditional logit model in several ways. Our primary innovations are to account for choice set variability created by selective admissions, to control for fixed unobserved differences between schools, and to permit greater preference heterogeneity which generates realistic substitution patterns between colleges. In this section, we review the basic setup of the model, discuss some critical issues involving identification and interpretation of the parameter estimates, and detail our estimation strategy.

4.1 Basic Setup

If the random components ε_{ij} in equation (1) are assumed to be independent and identically distributed across individuals and choices with the extreme value distribution, the probability that individual *i* is observed choosing college *j* is given by the simple conditional logit formula:

$$\Pr(Enroll_{ij}=1)=\frac{\exp(\delta_{ij})}{\sum_{k=1}^{J}\exp(\delta_{ik})} \qquad (2)$$

where $\delta_{ij} \equiv -\alpha_{1i}T_{ij}+\alpha_{2i}A_j+\alpha_{3i}C_j+\alpha_{4i}D_{ij}$ is the value function for school *j* as perceived by individual *i*. Note that student characteristics that do not vary across their choices (e.g. income or race) cannot enter independently into this basic model. In the base model, preference parameters do not vary across students: $\alpha_{ki}=\overline{\alpha_k}$. The coefficients $\overline{\alpha_k}$ parameterize the average preference for attribute *k* in the population. In a cross-sectional sample, the parameters of equation (2) are identified by differences in the enrollment shares across institutions and subgroups that are related to the variables of interest. If students value instructional expenditure, for example, then schools with more spending on instruction should have a greater share of all postsecondary students than schools with less spending. Coefficients on attributes that vary across students within schools will additionally be identified by within-school variation. For example, students facing a higher price for a given school (e.g. out-of-state students) should be less likely to attend if cost is a deterrent to enrollment. Unlike the multinomial logit of Manski and Wise (1983), this conditional logit model takes advantage of match-specific attributes between students and colleges for identification.

This model has at least four limitations. First, any component of unobserved demand ε_{ij} that is correlated with the included covariates will bias estimates of students' willingness to pay. This is the classic omitted variable bias problem. Second, selective admissions will effectively limit students' choice set to less selective schools. Attributes of less selective schools will thus appear more favorable, since more students will attend them. This is a specific form of omitted variable bias caused by a misspecification of some students' choice set. The consequence is estimated parameters will cofound school selectivity with student preferences. Third, the basic conditional logit model predicts that the substitutability of a pair of colleges is proportional to their initial enrollment shares, which is unrealistic if students tend to substitute between colleges with similar characteristics.[8] We now describe our strategy for addressing all three of these limitations.

4.2 Addressing Omitted Variable Bias

As with ordinary least squares, if observed college characteristics are related to unobserved (or un-controlled-for) college characteristics that also influence demand, then simple estimates of (2) may suffer from omitted variable bias. Total capacity is one possible confounder. For example, if very large schools have lower tuition or weaker academic standards, a choice model that does not control for size will tend to understate student willingness to pay for academic quality and overstate the disutility associated with high tuition. For this reason, many college choice models include some measure of enrollment. However, the inclusion of college enrollment makes it impossible to identify preferences for college characteristics that are invariant across students using a single cross-section.

To see this, imagine that one had data on all students and colleges in the population. In this case, the average likelihood of any particular student attending a particular school would be exactly that college's share of the college market. Hence, inclusion of fully flexible enrollment variables alone would lead to a perfectly fit model. If one estimates a model using a subset of the population, sampling variability alone will determine which schools have sample shares greater or less than what would be predicted by the college's (population-based) enrollment.

This complication arising from controlling for enrollment are similar to the issues that arise when including fixed effects to control for any unobserved desirable characteristics that may be related to a regressor of interest. Coefficients on variables that vary only at the school-level are not identified when enrollment or fixed effects are included. Importantly, this is not the case for college characteristics that vary across students within an institution such as price or distance. The coefficients on these variables are identified from differences in the likelihood of attendance among students with different values of the characteristic. For example, the coefficient on distance is identified by differences in enrollment shares among individuals living

[8] Differentiating shows that the marginal effect of a change in some attribute of college *j*, z_j, on the probability that college *j* is chosen is $dp_j / dz_j = p_j [1 - p_j] \alpha_z$ and the effect from a change for college *k* is $dp_j / dz_k = -p_j p_k \alpha_z$.

closer to or farther away from a given institution.[9] Coefficients on interactions between student and school characteristics (the α_{kx}'s) are identified in the presence of enrollment controls or fixed effects in a similar manner.

Much of the existing college-choice literature does not address this identification concern. For example, Long (2004) includes a quadratic in student enrollment to control for school size, and interprets the coefficients on instructional expenditures at the institution as measuring the demand for this attribute. Several structural models of college-choice do address this concern by fully specifying the underlying application, admission, and enrollment process and observing the distribution of student attributes across colleges (Arcidiacano 2005; Epple, Romano, and Sieg 2006).[10]

In order to identify the importance of student-invariant college attributes, we stack data from multiple cohorts and exploit variation in attributes and enrollment within schools across cohorts. If students are willing to pay for an attribute, schools with increasing levels of this attribute should see their enrollment increasing over time and one should observe schools with high values of this attribute entering the market. Our preferred model includes school fixed effects for the roughly 1300 colleges in our analysis sample, estimated through an iterative procedure in the spirit of Berry (1994) and Guimarães and Portugal (2009).[11] As a point of comparison, we also present models where we control for contemporaneous log enrollment. Avery, Glickman, Hoxby, and Metrick (2005) also incorporate fixed effects in a model of college choice and use these fixed effects estimates to construct revealed preference college rankings.

In a model with college fixed effects, our identifying assumption is that changes in

[9] Similarly, the in-state versus out-state tuition difference helps identify the coefficient on price by a comparison of the likelihood of in-state versus out-state students attending a particular college. One limitation is that many public universities place a cap on the number of out-of-states students they enroll, which may be correlated with in-/out-of-state tuition differentials.

[10] For instance, high-ability students (with many options) are revealing their preference for a particular college's characteristics when they decide to go there. This strategy requires that one make some assumption about the college admission process. If, for example, one were willing to assume that colleges select the most academically talented students that apply, then one could conclude that, all else equal, a college with higher-ability students enjoys higher demand than a similarly-sized school with lower-ability students. And, then one could infer that students had a preference for specific attributes of this institution. The revealed preferences approach to ranking institutions, put forth by Avery et. al (2005), is also related.

[11] Briefly, the approach iterates between estimating the main model parameters assuming a given set of fixed effects, then updating the fixed effects to equate predicted and sample probabilities. Standard errors are found by inverting the numerical hessian for the entire coefficient vector (including the fixed effects).

college attributes are uncorrelated with changes in unobserved tastes for individual colleges. For instance, if colleges that increase spending on student services also strengthen other favorable attributes (desirable alumni network), then our estimates will overstate the causal effect of increases in student services on colleges' ability to attract students. Similarly, this model implicitly assumes that college characteristics are exogenous from the perspective of school administrators. Colleges clearly have some discretion over characteristics such as amenities and tuition and could alter them in anticipation of (or in response to) demand changes, creating an endogeneity problem. While we cannot rule out this possibility, we believe that the potential bias introduced is minimal.[12]

In order to control for some potentially important time-varying factors, in some specifications we also several other time-varying characteristics associated with each college. For example, we control for the unemployment rate in the state in which a college is located in the year in which the cohort would have been applying to college in order to account for the fact that students may be reluctant to attend college in an economically depressed area if they intend to reside in the area after graduation. In some specifications, we control for binary indicators of whether the college is located in the same state and/or region in which the student attended high school. This is meant to control for hard-to-observe factors such as family connections that will influence a student's college choice beyond the distance and cost variables that we already have in the model and that will pick up several key differences between in-state and out-of-state schools. Lastly, in some specifications we replace college sticker price with the estimated net price to account for price discounting across students and schools which may confound estimates of demand preferences.[13]

Finally, there are several other limitations to the panel model described above. While colleges have some flexibility to adjust enrollment and tuition, neither of these factors is perfectly elastic (in the short-run). For example, an individual college could not quadruple the

[12] It should be noted that if the market responds to a demand for college amenities with the creation of new amenity rich schools, then the inclusion of school fixed effects would tend to understate the value students place on amenities. In practice, the entry and exit of colleges seems unlikely to be important in our analysis. Of the 2,853 college-years in our sample of "regular" four-year colleges, 46 were open only in 1992, 97 were open only in 2004 and the remaining 2,710 were open in both years. When we limit our sample to the 2,458 college-years that were ever selected by individuals in our student-level data, 13 were only open in 1992 and 51 were only open in 2004.

[13] To implement this, we estimated a model with the net price ratio (price minus grants over price) as the dependent variable using the 1996 and 2004 National Postsecondary Student Aid Study. The model was estimated separately for six groups (defined by race X sector X in-state) separately by year and with many interactions. The estimated model was used to predict net price for students in our analysis sample. Results are described in Appendix E.

size of its incoming class to accommodate increased demand due to short-run constraints in physical capital. Similarly, there are probably at least some barriers to entry in the college market. These frictions will lead us to understate student preferences for college characteristics in the model.

4.3 Admissions Selectivity and Unobservable Choice Set Variation

A second concern with the basic conditional logit model is that selective admissions necessarily prevents some people from attending certain schools, even if they desire to do so. The standard approach models the enrollment choice out of a set of potential schools, which may include many schools to which the student did not apply or to which the student would not have been admitted had s/he applied. In doing so, this approach confounds the enrollment and admissions decisions and may lead to biased estimates of student preferences. More generally, many discrete choice settings are characterized by variation in the effective choice set faced by decision-makers, which is often unobserved and/or endogenous. The IO demand estimation literature has primarily focused on settings where all products are generally available to all consumers and paid little attention to settings with considerable choice-set variability. [14] Choice set variability is pervasive in many situations beyond education, including choice of residence, job or occupation, and products that experience supply constraints and stock-outs.

Conlon and Mortimer (2010) describe two sources of bias in demand estimates when failing to account for unobserved choice set variation (product stock-outs in their case). First, demand estimates will be censored: products that sell out will have actual demand that is higher (at observed prices) than sales would suggest. In our context, preferences for attributes possessed by desirable (but supply-constrained) schools will be underestimated. Second, forced substitution to less desirable (but unconstrained) products (schools) may overstate the importance of unconstrained products' attributes. These biases will cloud our understanding of the relative importance of various attributes in college choice, causing schools to appear to be greater substitutes than they actually are, and undermine the credibility of various policy simulations, such as the likely effect of new college opening or demand responses to changes in college characteristics. Conlon and Mortimer (2010) find substantial bias in vending machine product demand estimates that fail to account for stock-outs. Monte Carlo evidence by Desposato (2005)

[14] Seminal papers by Berry, Levinsohn, and Pakes (1995, 2004) and Nevo (2001) focus on cars and breakfast cereal, respectively.

also finds that choice set variability can bias standard conditional logit estimates.

Suppose the set of all possible choice sets that can be formed by the J colleges is denoted by C* and individual *i*'s choice set is denoted by C_i. The unrestricted choice set (all schools are available) is given by C^J. Then the unconditional probability of the enrollment outcome we observe for student *i* and school *j* in the data is the probability of *i's* enrollment choice conditional on a particular choice set and the likelihood of this being their individual choice set, integrated over all the possible choice sets.

$$\Pr(Enroll_{ij}) = \sum_{C_i \in C^*} \Pr(Enroll_{ij} \mid C_i) \Pr(C_i \mid C^*) \quad (3)$$

where $\Pr(Enroll_{ij} \mid C_i) = \dfrac{\exp(\delta_{ij})}{\sum_{k \in C_i} \exp(\delta_{ik})}$. If all individuals could choose from all schools, then $C_i = C^J$ for all *i* and this reduces to equation (2).

There are a number of ways that have been proposed to address this issue. First, one can ex-ante specify the choice set for each individual directly. While easy to implement, this approach inevitably causes errors: some alternatives that are excluded from the choice set may be chosen. A second approach is to control for characteristics that may determine choice set variation. In this vein, Long (2004) includes flexible interactions between a college's academic quality and student ability (measured by test scores) to control for the likelihood that an individual would have been admitted to the school. A conceptual limitation of this approach is that it does not allow one to distinguish between admissions constraints and heterogeneity in references by student ability. In addition, the inclusion of these covariates may do a poor job of approximating the very non-linear constraint on choice imposed by selective admissions.[15]

A third possibility is to estimate a model of choice set determination explicitly. This is the approach taken by Arcidiacono (2003) and advocated by Horowitz (1990), which is easiest to implement when the choice set is actually observed. In our setting, the choice set is partially unobserved since we do not know the full set of schools applied and admitted to (for 1992) and do not know admissions outcomes for schools not applied to. In this case, estimation treats the individual-specific choice sets as an unobserved variable that is integrated out of the likelihood

15 Through Monte-Carlo simulations, Desposato (2005) finds that controlling for covariates thought to determine choice set selection does a poor job of mitigating bias.

value. To implement this, we would approximate (3) through simulation by drawing a number of different choice sets, calculating the probability of enrollment given this choice set, then averaging across all the repetitions. Choice sets would be sampled from their probability distribution, which is implied by an estimated model of admissions.[16] For a large enough number of replications, this should approximate the true unconditional likelihood of enrollment. This is in the spirit of the approach taken by Conlon and Mortimer (2010) and advocated by Desposato (2005).

However, a simulation-based estimator is computationally intractable for our models that include school fixed effects because the simulated likelihood must be calculated for each iteration of the parameters and fixed effects. Instead, we implement a computationally-tractable alternative to this simulation-based approach that is simply a standard conditional logit where each alternative's value function is weighted by the estimated probability of inclusion in the choice set,

$$\Pr(Enroll_{ij}) = \frac{\psi_{ij}\exp\left(\delta_{ij}\right)}{\sum \psi_{ik}\exp\left(\delta_{ik}\right)} \tag{4}$$

In our context, the weight ψ_{ij} is simply the predicted probability that individual *i* would be admitted to school *j* if he or she applied.

$$\psi_{ij} = pr(j \in C_i) \tag{5}$$

If the number of possible schools is sufficiently large, (4) will provide a good approximation of the simulation-based likelihood described above.[17] This weighted model is mathematically

[16] This would involve six steps: (1) Estimate admission as a very flexible function of school and individual characteristics and their interactions on the sample of schools students applied to. (2) Draw a set of admissions shocks for all individual-school observations. This is a vector of length NxJ (where N is number of individuals and J is number of schools). (3) Given the parameters of the admissions model and the vector of admissions shocks, simulate admission for all individual-school observations. These admissions outcomes determine the choice set for individual *i* during replication *r*. (4) For each observation, calculate the probability of enrollment given this choice set and store the vector of enrollment probabilities. Predicted probabilities come from the conditional logit model and the predicted probability is zero for schools not in the choice set. (5) Repeat steps 2-4 for R replications. (6) Average the predicted enrollment probabilities across the R replications to approximate the unconditional probability of enrollment.

[17] Appendix D presents simulation results which suggest this approximation is good. Using the parameter values contained in column (2) of Table 5 to estimate choice probabilities, the correlation between the observation-level likelihood implied by the weighted and simulation-based approaches is 0.9879 overall, with the approximation being better for individual-school observations with a high likelihood of acceptance. Table D1 shows these correlations separately by the predicted probability of admission.

equivalent to including $\ln(\psi_{ij})$ as a covariate whose coefficient is constrained to equal one and can easily be estimated in standard statistical packages.

To approximate ψ_{ij} , we use the predicted probability of acceptance from a probit model that is estimated with data on college admission outcomes for students and colleges in our sample using a very flexible function of student and school characteristics and their interactions.[18] Appendix C provides more details on the admission model sample and estimates.

The critical identifying assumption in our approach is that, conditional on the detailed set of student and school characteristics we include in the models, there are no unobservable factors that are simultaneously correlated with the likelihood of admissions and enrollment.

4.3 Relaxing the IIA Assumption

A well-known limitation of the standard conditional logit model is the restriction it places on the error terms. While our preferred specifications permit tastes to vary with observed student attributes such as academic ability and socioeconomic status, the conditional logic model is not able to accommodate tastes that vary with unobserved variables or purely randomly. For instance, if tastes vary with respect to an unobserved variable, then it errors are necessarily correlated over alternatives and its variance also varies over alternatives (Train, 2003) and the logistic model is misspecified. Thus the standard conditional logistic model imposes the property of independence from irrelevant alternatives (IIA). That is, the relative choice probabilities for any two alternatives will not depend on the presence or characteristics of any other alternatives. The relative likelihood of choosing one specific college over another is the same regardless of the other colleges available.

One implication is that cross-elasticities will exhibit proportional substitution. Since the ratio of probabilities between two alternatives is always the same, any change in the characteristics of a third alternative will impact the two alternatives by the same proportion. For instance, the conditional logit model predicts that if Cal State Long Beach increased instructional spending, then the share of students attending Cal State Northridge and Harvard University would decrease by proportionately the same amount. This pattern of substitution seems unrealistic.

[18] Student and school characteristics include student race, gender, SES, high school GPA and standardized achievement scores along with measures of the school's selectivity such as the average SAT score of students in the school. Admissions models are estimated separately by the triple interaction of race, college sector, and in-state status.

To address these concerns, our preferred model lets preference parameters vary with observable student characteristics: $\alpha_{ki} = \overline{\alpha_k} + \alpha_{kx} X_i$. The coefficients $\overline{\alpha_k}$ parameterize the average preference for attribute *k* in the population and the coefficients α_{kx} captures how preferences for attribute *k* vary with individual characteristics such as gender, ability and socioeconomic status.[19]

4.4 Remaining Threats to Identification

Though we have addressed a number of limitations in the previous literature, our strategy still has a number of potential threats to identification. First, like most fixed effect panel data models, we assume that changes in our variables of interest are uncorrelated with changes in unobserved determinants of demand. For instance, if colleges that increased their spending on consumption amenities also increased spending on marketing, then we may be attributing the effect of the marketing campaign to the importance of spending on consumption amenities.[20]

Second, variables we interpret as "consumption" may actually measure something that provide labor market returns, and thus be properly categorized as "investment." While the correspondence between our measures of consumption and students' self-reported college attributes give credence to our interpretation, we cannot entirely rule this out. Though this would not necessarily invalidate the credibility of our estimates of the effect of these variables on student choice, it would change their interpretation.

Third, our models assume that students are informed about college characteristics. If some information is hidden, we might misinterpret a lack of demand for an attribute with a lack of information about the attribute.

Lastly, our estimates of the heterogeneity of preferences by student characteristics depends on our ability to adequately predict students' probability of admission to each college. Unobserved student characteristics related to both admissions and enrollment may introduce bias in our estimates of this preference heterogeneity, which may conflate preferences with selective admissions.

[19] Unobserved heterogeneity is not yet incorporated in the model, though future versions of this paper will do this as well using a mixed logit model (see Train 2009). The current draft assumes that there is no preference heterogeneity within groups defined by these demographic variables. If these demographic variables capture a sufficient amount of the variation in preferences then the substitution patterns will be realistic.

[20] To address this identification concern we are currently exploring several instrumental variables and control function strategies.

5. Data

In our analysis, we combine student-level data from two nationally representative cohorts of high school seniors with college-level data on approximately all four-year colleges in the U.S. This section briefly describes the key features of the data used, including the sample construction. For additional detail, see Appendix A.

5.1 College-Level Data

We combine data from a number of different sources to construct an unbalanced panel dataset of postsecondary institutions for 1992 and 2004. We limit our sample in several ways to facilitate our focus on amenities arguably related to direct, immediate consumption value. First, we limit our sample to public and non-profit private undergraduate four-year schools only, excluding all two-year (or less) schools, all for-profit schools, and schools offering professional degrees only. Second, we drop specialized divinity, law, medical, specialized health (e.g. nursing), and art colleges, though we keep engineering, teaching, military, and business colleges. Finally, we drop schools with an average of fewer than 50 freshmen or 300 FTEs during our analysis years in an effort to eliminate remaining specialized schools which are arguably not in many students' consideration set.

Total undergraduate tuition and fees for in- and out-of-state students were obtained from the IPEDS Institutional Characteristics surveys, as were sector (public or private), and level (4-year or 2-year). From this source we also obtained information on religious affiliation, same-sex status, historically black college or tribal college status, and whether the institution is focused on a specific major area (business, engineering, education, health, law, seminary, etc). Total freshmen enrollment, freshmen enrollment by state, and full-time equivalent students (including undergraduate and graduate students) were obtained from the IPEDS Fall Enrollment surveys.

We use institutional spending in various categories as our primary measures of academic quality and consumption amenities. We use expenditures on instruction and academic support per FTE as a measure of the institution's academic quality. The expenditure data comes from the IPEDS Finance survey and the Delta Cost Project.[21] These categories include expenses for all forms of instruction (i.e., academic, occupational, vocational, adult basic education and extension sessions, credit and non-credit) as well as spending on libraries, museums, galleries, etc. Following the prior literature, in most specifications we also use the average SAT score of

[21] This survey was changed considerably in 2000, but the spending categories are mostly comparable across year.

students in the college as a measure of academic quality. We obtained the average SAT percentile score (or ACT equivalent) of the incoming student body from Cass Barron's *Profiles of American Colleges* (1992).[22] For 2004, we used the average of the 25th and 75th SAT percentile, which we obtained from IPEDS.

Longitudinal data on consumption amenities are more difficult to come by. Our primary measure of consumption amenities is spending on student services and auxiliary enterprises. Spending on student services includes spending on admissions, registrar, student records, student activities, cultural events, student newspapers, intramural athletics, and student organizations. Auxiliary expenditures include those for residence halls, food services, student health services, intercollegiate athletics, college unions and college stores. All spending measures have been deflated by the CPI-U and are in 2009 dollars.

Finally, in some specifications (not reported, but available from the authors) we control for the cost-of-living in the geographic areas in which each college is located. The cost-of-living index is based on the cost of a weighted bundle of consumer goods. The data is collected annually for a variety of cities across the United States by the Council for Community and Economic Research and local Chambers of Commerce.

Table 1 presents summary statistics of the college data, separately by sector for 1992 and 2004. Real tuition costs and spending on instruction and student services increased considerably during the 1990s, though there are differences across sectors. Public institutions saw a greater proportionate increase in tuition prices, while private institutions saw larger relative increases in both forms of spending. Furthermore, the average SAT percentile score of colleges' students actually declined over this period.

Many of these measures are highly positively correlated, as depicted in Tables 2.[23] Log per-student spending on instruction and student services/auxiliary are positively correlated with each other (correlation coefficient 0.55-0.60) and with tuition. Schools that have high SAT-scoring students tend to spend more on both instruction and student services and also charge higher tuition. Cross-sectional correlations are very similar for 1992 and 2004.

Most important for our identification of models with fixed school effects is the presence of independent variation in our main school characteristics over time. For example, schools must

[22] We thank Bridget Terry Long for providing us this data, which she used in her 2004 paper (Long 2004).
[23] Correlations that weight by enrollment yield comparable results.

have changes in spending on instruction that are independent from changes in tuition or spending on services. The bottom panel of Table 2 reports correlation coefficients on changes in college characteristics from 1992 to 2004. In comparison to the cross-section, the correlations between changes in variables over time is much smaller.

5.2 Student-Level Data

We combine two nationally representative samples of the high school classes of 1992 (National Educational Longitudinal Study, NELS) and 2004 (Educational Longitudinal Survey, ELS). Prior work by Long (2004) has utilized data from two earlier cohorts, the high school classes of 1972 (National Longitudinal Survey, NLS72) and 1980/82 (High School and Beyond, HSB82). We exclude these from our analysis because they do not have sufficient information on college applications/admissions, which is necessary to properly account for separating admission from enrollment decisions.[24]

These longitudinal surveys follow students from high school into college. We limit our sample to individuals who graduated from high school, attended a four-year institution within two years of expected high school graduation, attended a college in our choice set, and were not missing key covariates (test scores, race, gender, family SES, college choice, etc).

We assign out-of-state tuition levels to individuals residing in all states other than the one in which the institution is located, so (at this point) we do not take into account tuition reciprocity agreements between neighboring states. Tuition does not vary by in-state status for private institutions. As a proxy for the distance between a student's home and a college, we calculate the distance between the centroid of the zipcode in which the student's high school is located and the centroid of the zipcode in which each institution is located.

Table 3 presents summary statistics for our analysis sample. The bottom panel presents statistics on the colleges attended by our sample. Over our analysis period, the real cost of tuition increased almost forty percent, from $8,864 in 1992 to $12,295 in 2004, while the average distance traveled to college increased from 190 to 208 miles. Schools attended by our sample increased spending on instruction 17 percent over the period and spending on student services by roughly 12 percent. Interestingly, the fraction attending a religious school remained roughly

[24] The NLS72 actually does include a short list of several schools to which each student applied, but in over 90 percent of cases, the student indicates that he or she was accepted to the school. Hence, this information does not provide sufficient variation to allow us to estimate a credible admissions model.

constant at around 18 percent, as did the fraction attending a single-sex school (roughly 3 percent) and the fraction attending a historically Black college (5 percent).

Each of these surveys asked high school seniors what factors they viewed as most important in selecting a college. These self-reported preferences provide some interesting descriptive information, and allow us to validate some of our more objective college characteristics (see below). The bottom panel of Table 4 shows that the fraction of students citing the reputation of the college, the courses available and the availability of financial aid as "very important" has increased from 1992-2004. The fraction citing factors such as athletics and social life also increase substantially from 1992 to 2004.

For the purpose of the analysis below, we create three composite measures based on a simple average of these items. The variables, standardized using the mean and standard deviation from the 1972 cohort, capture the self-reported value that students place on academics, cost and social life.[25] The summary statistics for these composites shown in Table 3 are for the analysis sample, and show an increasing value placed on all three factors. In the analysis below, we rely primarily on the across-student variation in these measures rather than the across-cohort variation.

6. Results

6.1 Extensions to Previous Work

To provide a direct comparison with previous work, we first extend the analysis of Long (2004) by including measures of college consumption amenities into her conditional logit specifications. These results, reported in Appendix B, indicate that we are able to replicate her main findings and, more importantly, several different measures of college "consumption amenities" are significant predictors of student choice above and beyond all of academic measures included in Long (2004). Furthermore, the inclusion of these measures diminishes the estimated importance of instructional expenditure.

Table 4 presents results from comparable models estimated separately by cohort, but using a specification that mirrors that used in our subsequent analysis. Table 4 shows the odds ratios and standard errors from our conditional logit model described above, separately for the

[25] This normalization reflects our use of the 1972 cohort in earlier analysis. Subsequent analysis will normalize this measure using the 1992 cohort. The normalization base will not have any effect on our results.

1992 and 2004 cohort. Following the prior literature, we include log enrollment to control for school size. Given the cross-sectional identification concerns raised in the previous section, we do not interpret these specifications as providing good estimates of preference parameters. Rather, this analysis provides a benchmark for subsequent analysis and illustrates the importance of including consumption amenities and accounting for selective admissions when estimating college choice models.

In columns 1 and 5, we see that tuition and distance are negatively associated with student choice while enrollment, instructional spending and mean SAT score are positively associated with choice. In columns 2 and 6, we add the log of per pupil spending on student services and auxiliary enterprises, which we argue measures the level of consumption amenities at the college. Conditional on the measures of cost, distance and academic quality, we see that spending on student services is a significant predictor of enrollment. Specifically, the odds ratio of 2.45 in column (2) indicates that a doubling of spending on student services is associated with a 145 percent increase in the likelihood a student will attend a given school in 1992. Note that the magnitude of this effect is even larger than the effect of instructional spending. Because school mean SAT and instructional spending are arguably closely related proxies for academic quality, in columns 3 and 7 we show results for a model that excludes school mean SAT. The odds ratio on instructional spending does increase, but spending on student services still remains a stronger predictor of enrollment than instructional spending.

As noted earlier, however, failing to account for whether a student would be accepted to a given school may bias estimates of the importance of college attributes. To account for selective admissions, the models shown in columns 4 and 8 weight each student-college observation by the predicted probability that the student would have been accepted to the college. As expected, the coefficients on both measures of academic quality – instructional spending and mean SAT – increase considerably. Failing to adequately account for selective admissions may bias estimates of students' preferences for college attributes that are also related to admissions difficulty.

6.2 Preference Estimates: No Heterogeneity

Table 5 provides estimates of the choice model pooling the 1992 and 2004 cohorts and imposing homogeneity in student preferences. The first two specifications do not include college fixed effects and demonstrate patterns mirroring those for the individual years in Table 4. Cost and distance are major predictors of where students choose to enroll, as is spending on student services and instruction. Failing to account for admissions difficulty will cause estimates of the desirability academic characteristics – instructional spending and mean SAT – to be understated. Estimates in column (2) imply that a doubling of spending on student services is associated with an 85 percent increase in the likelihood of enrolling. The increase in enrollment probability associated with a comparable increase in instruction is 51 percent. In order to help interpret the magnitude of these results and to quantify the relative tradeoffs that students are making, the bottom panel of the table reports measures of "willingness-to-pay" (WTP) for each college attribute. For each college attribute, the WTP is given by the (negative) ratio of the estimated coefficient on that attribute to the estimated coefficient on log(cost).[26] For example, the WTP of .214 for instructional spending in the bottom panel of column 2 indicates that students are willing to pay roughly 0.21 percent more to attend a school that spends 1 percent more on instruction per student. In contrast, the student would be willing to pay about 0.32 percent more to attend a school that spends 1 percent more on student services. The WTP of .011 on school mean SAT indicates that a student would pay 1.1 percent more to attend a school whose mean SAT score is 1 percentage point higher on the national distribution. In order to attend a top quartile school (in terms of mean SAT measure) instead of a bottom quartile school, a student would be willing to pay nearly 50 percent more (i.e., .011 x (79-34) = .495). The -0.573 WTP for distance indicates that a student would be willing to pay 0.57 percent more to attend a school that was 1 percent closer.

As noted earlier, student-invariant college attributes can be identified in a model that relies on cross-sectional variation only if one is willing to assume that simple controls for school enrollment are sufficient to control for unobserved characteristics that are correlated with size and desirability of the college. Specification (3) includes college fixed effects, meaning that identification comes from within-college changes over time in attributes that are associated with within-college chances in enrollment. The inclusion of college fixed effects changes the results in several important ways. First, the importance of cost increases noticeably, with the odds ratio going from .146 in column 2 to .008 in column 3. This suggests that expensive colleges also possess unobservable qualities that are attractive to students. Not accounting for these fixed unobservable attributes may cause cost to appear to be a less important consideration than it truly

[26] Standard errors on the WTP measures are calculated using the Delta Method.

is. This is a finding that is common in the differentiated products literature: accounting for unobserved product characteristics typically makes the effect of price more negative. The coefficient on distance does not change much when fixed effects are included.

In contrast, the coefficients on the other college attributes decline substantially. The coefficient on instructional spending actually becomes negative, but is significant only at the 10 percent level. The coefficient on our measure of consumption amenities declines as well (odds ratio = 1.38), but remains statistically significant and suggests that students are willing to pay 6.6 percent more for a 100 percent increase in spending on student services. The odds ratio on school mean SAT drops from 1.021 to 1.005 and is significant only at the 10 percent level. The WTP of .001 for school mean SAT in column 3 indicates that students would pay roughly 0.1 percent more for a 1 percentage point increase in school mean SAT. In order to attend a top quartile school (in terms of this mean SAT measure) instead of a bottom quartile school, a student would be willing to pay roughly 4.7 percent more (i.e., .0010 x (79-34) = .0466).

Columns (4) and (5) include controls for other regional and geographic characteristics that may be correlated with college amenities. Column (4) includes controls for the annual unemployment rate in the institution's state and log of the number of high school graduates, to control for state-specific labor market conditions and cohort crowding, both of which may influence college choice. Neither of these controls is significant or changes the results. Column (5) includes indicators for whether a college is in the student's home state and region, to account for non-linearities in preferences for proximity coupled with changes in the geographic distribution of students over time. Both are strong predictors of college choice, but their inclusion does not qualitatively change the estimated importance of the other college amenities.[27]

Financial aid, which reduces the actual net price students pay, is another source of omitted variable bias. If schools that offer more generous financial aid packages are also those with higher cost or greater spending, then our base estimates will understate the importance of the former and overstate the importance of the latter. Specification (6) addresses this concern by using the log of predicted net price (tuition and fees plus room & board minus total grant aid) as

[27] We also estimated specifications that included controls for the cost of living (normalized within year) in each college's city or town, to absorb variation in spending due to higher prices which may not reflect differences in real amenities. Again, estimates of students' willingness to pay for various college amenities are unaffected by this control. To address multi-colinearity concerns with including two distinct measures of academic quality (instructional spending per student and average SAT scores), we also estimated specifications that exclude average SAT score. This has virtually no impact on the other estimates and instructional expenditure remains insignificant. These results are available from the authors upon request.

the measure of cost. The importance of school mean SAT increases when net price is accounted for, as expected, but the importance of price to student decisions is actually diminished. Comparing columns (3) and (6), the estimated preferences for services and instruction spending is not changed much when net price is accounted for.

6.3 Preference Estimates: Heterogeneity

The results presented above suggest that, on average, students value institutions' spending on consumption attributes more than spending on instruction. However, preferences are likely to differ between students for many reasons and this preference heterogeneity has implications for policy and schools. Preference heterogeneity will impact the elasticies that colleges face in response to changes in their characteristics. Failing to adequately capture preference heterogeneity thus has the potential to mischaracterize the demand-side incentives colleges face.

To examine how preferences for college attributes vary with observable student characteristics, we permit student preferences for college attributes to vary with sex, student ability and family income. Table 6 reports coefficient estimates for models that include interactions between these three student characteristics and the five college attributes (odds ratios are difficult to interpret with many interactions, so raw coefficients are presented).

The first specification accounts for selective admissions, but does not include college fixed effects. Column (2) includes college fixed effects and column (3) includes fixed effects and uses predicted net price as the measure of college cost. Across all three specifications, heterogeneity is substantial. Wealthier students (higher SES) are substantially less sensitive to price and distance and higher achieving students are less sensitive to distance. Male students are more price sensitive than female students.

High-ability students have a much greater preference for academic quality, both in the form of instructional spending and mean SAT. Interestingly, this pattern changes little when school fixed effects are included or when net price is used as the measure of cost. Recall that these models account for the predicted probability of acceptance that incorporate the 12^{th} grade test scores along with other measures of academic aptitude so this finding is not simply an indication of the greater likelihood of acceptance to elite institutions among such students. A similar pattern is observed in the preference for school mean SAT, with preference for this characteristic increasing with both ability and income.

Differences in valuation for consumption amenities by student ability and income is much less pronounced, through the net price specification (column (3)) does suggest that higher income students have a greater preference for consumption amenities while higher achieving students place less value on this. Figures 1 to 5 summarize the variation in predicted WTP across our sample, where WTP is predicted using the model estimates for specification (2). Figure 1 plots the overall distribution of WTP for each college attribute, demonstrating that there is substantial predicted heterogeneity in students' willingness to pay for all college characteristics. The WTP for student services and for SAT is positive for almost all members of the sample, yet the same is only true for instruction for a limited number of individuals. Figures 2 through 4 present comparable distributions for certain sub-populations to better quantify the importance of student characteristics on preferences. These graphs demonstrate a strong variation in preference for academic quality associated with academic preparation. Very high achieving students tend to derive greater value from high academic quality. In fact, the distribution of estimated preferences for instructional spending does not overlap between students in the top and lowest test score terciles. SES does contribute to heterogeneity, particularly on the WTP for student services. Preference variation by sex is minimal.

Figure 5 shows mean WTP for nine subgroups defined by sex, test scores, and SES. There is substantial variation in the value placed on consumption and instructional spending, with the greatest willingness-to-pay for student services being low-ability, high-income students. Instructional spending only has a positive WTP for high-ability, high-SES students.[28] Income is a strong predictor of preferences, even within ability categories.

6.4 Interpretation as "Consumption Amenities"

We have documented a substantial enrollment response to spending on student services and auxiliary enterprises, which we interpret as evidence of the importance of consumption considerations in students' decisions. Evidence in favor of this interpretation is presented in Table 7. In this analysis, we estimate a model that includes interactions between our five college attributes and the three self-reported student "preference" measures described earlier. Recall that these measures are standardized composite variables that reflect how the students, as 12th graders, reported the importance of different college characteristics in their decision to enroll in a

[28] In results not reported here, we also find that students are substantially more likely to attend institutions that match their background (e.g., Black students attending historically Black colleges, Catholic students attending Catholic colleges, etc.).

particular school. We view this specification as a useful check on the validity of our college attribute measures. For example, if spending on student services were really capturing something about the consumption value of an institution, we would expect students who report that a school's social life is important to be more likely to attend these institutions. Similarly, if instructional spending were a good proxy for academic quality, students who report academics to be very important to them should be more likely to attend schools with higher spending on instruction. Indeed, we find exactly these patterns, bolstering our confidence in the college attribute measures we use.[29] Furthermore, these estimates control for selective admissions so these patterns do not simply reflect differences in the likelihood of acceptance to schools with different characteristics between students reporting "social" vs. "academic" factors as being important to their decisions.

7. Implications for Colleges

We now use our estimates of the college demand model to characterize the consequences of student preferences, and their variation across students and between attributes, for institutions. We use the estimated conditional logit model to simulate changes in patterns of demand if colleges were to alter their characteristics in an attempt to attract students. We took each individual college and altered a single characteristic one at a time, while holding all other characteristics of it and of all other colleges constant. Then we recorded how the entire pattern of enrollment across all colleges changed. For each individual college, we examined responses to four changes:

- 1% increase in cost;
- 1% increase in spending on services;
- 1% increase in spending on instruction; and
- 1 percentile point increase in mean institution SAT.

The simulated change in enrollment provides a guide to the demand-induced incentives colleges have to alter their characteristics. Since colleges may care about different attributes of their students, we documented the enrollment response of three student groups: all students, high SES

[29] As an additional test, we examined the relationship between our college attributes and the subjective assessments of college "quality of life" and "quality of academics" presented in the Princeton Review guidebooks. Students attending colleges with more spending on student services rate the quality of life of the institution much higher, whereas instructional spending has little correlation with subjective quality of life. By contrast, students rate colleges with high instructional expenditure or higher student services expenditure as having a better "academic environment."

students (those above the 75th percentile), and high achieving students (those above the 75th percentile of math test score). Given our analysis sample size of 1261 individual colleges, this results in 1,590,121 own- and cross-college enrollment responses to each of the four different policy changes.

Figure 6 plots the distribution of predicted own total enrollment elasticities with respect to each of the four college characteristics. The preferred model permits preference parameters to vary by sex, math scores, and SES, though the implied distribution of elasticities given the homogeneity preference model is shown for reference. Permitting preference heterogeneity for individuals results in much greater enrollment elasticity variation across schools since it permits the marginal student at each school to have different preferences and hence different responsiveness to institutional characteristics.[30]

Consider first the distribution of price elasticties shown in the top-left panel. The entire distribution of elasticities falls to the left of zero, indicating that all schools experience a downward sloping demand curve (i.e., a negative enrollment response to higher tuition). Overall demand is very price-elastic: the median price elasticity among colleges is -0.045, indicating that a 1% increase in tuition is associated with a 4.5% decrease in total enrollment. The panel in the top-right corner shows that all colleges are estimated to have a positive total enrollment response to marginal increases in student services spending. While most colleges are estimated to have a positive total enrollment response to marginal improvements in average SAT score, roughly 22 percent of institutions in our sample are estimated to have a negative elasticity of enrollment with respect to improvements in mean SAT score. Consistent with the results presented in Table 5, the vast majority of colleges appear to have a negative total enrollment response to increases in instructional spending.

Figure 7 plots the implied own-elasticities for enrollment of high SES and of high achieving students. However, each distribution still includes one elasticity per college. High achieving students are particularly responsive to improvements in academic quality, both in the form of average SAT and instructional spending. In fact, high achieving students are the only subgroup that responds positively to instructional spending; almost all colleges can attract more high achieving students by increasing instructional spending, though this usually comes at a cost

[30] Even with no preference heterogeneity, there is still limited variation in responsiveness across colleges due to differences in the distribution of distances and costs across students.

to their ability to attract other students. There is less variation in the responsiveness to student services spending across institutions. Most institutions can increase total or high SES enrollment by increasing student services, though the response of high-achieving students is smaller. The implication is that most colleges face a trade-off: increases in instructional spending will attract high achieving students, but may deter enrollment from a broader student body. Increases in service spending, however, will attract all types of students (though disproportionately lower-achieving students).

What predicts this variation in enrollment responsiveness across institutions? Do low SAT schools face a different demand response when they change their characteristics than their more selective peers? Figure 8 depicts the total enrollment own-elasticity with respect to the four college characteristics, by average student SAT score percentile at baseline. Though the own-price elasticity is similar across institutions with very different levels of selectivity, there are clear differences in responsiveness to other college characteristics. For instance, less selective schools (i.e., schools with low average student SAT scores) experience a greater enrollment response to a one percent change in student service spending than more selective schools. This is due, in part, to the fact that students on the margin of attending these institutions are lower ability students who are more sensitive to student service spending than their higher-ability peers. By contrast, the responsiveness of enrollment to changes in academic quality is higher at more selective schools. Students on the margin of attending more selective schools tend to place greater value on academic quality.

Figure 9 repeats this analysis for two important student subgroups: high achieving and high SES students. High achieving students are less responsive to service spending and more responsive to academic quality overall, but the difference in responsiveness across institutions of different selectivity levels is more muted. One implication is that institutions of very different selectivity face relatively similar incentives for attracting the most high-achieving students, but very different incentives when trying to attract students overall. The responsiveness of high SES students follows a similar pattern as students overall.

Figure 10 repeats this analysis, but using the estimates from the model with no student preference heterogeneity. This graph demonstrates how important it is to account for preference heterogeneity – without it the response to all characteristics appears to be similar across institutions and student groups. This pattern can be both quantitatively and even directionally incorrect since some colleges may face negative enrollment responses when they increase instruction or academic quality, while other colleges may see a positive response overall or for certain subgroups. Heterogeneity in institution response and incentives is masked without allowing for individual preference heterogeneity.

How do the actions of other institutions impact own-institution enrollment? Our simulations also reveal the cross-institution elasticities, such as how the enrollment at one school changes when another school raises its price or spends more on student services. While there are many ways one could examine these cross-elasticities, Figure 11 summarizes how enrollment responds to colleges within and across selectivity groups, operationalized by SAT quintiles. The four panels show the pattern of cross-elasticities for changes in four different college characteristics. Within each panel, the height of the bars represent the median percent enrollment change at schools in SAT quintile j when a single school in SAT quintile k changes a characteristic. For instance, the first bar in the top left panel represents the median enrollment response of colleges in the bottom SAT quintile to a reduction in cost at individual colleges in this same quintile. The next bar is the response at schools in the first quintile to an increase in cost at individual colleges in the second quintile (j=2).

Looking across the entire set of histograms, several interesting patterns are evident. First, colleges are generally more responsive to changes at schools with similar mean SAT scores. On Figure 11, this can be seen by comparing the height of bars for a given k group across the j group categories. Schools do have the largest impact on other schools in their same SAT group. Given earlier results that schools in different selectivity groups have different incentives for influencing enrollment, the relative high within-group cross-elasticities may reinforce these differences in incentives across groups.

Second, the cross-responses are asymmetric. Enrollment in colleges at all selectivity levels is responsive to changes in highly selective colleges, while enrollment at these more highly selective colleges are largely unaffected by changes in less selective colleges. This is particularly true in the market for high-achieving students, whose enrollment decisions are particularly sensitive to the characteristics of top-tier schools.

In results not reported, simulations that treat preferences as homogeneous will tend to understate the competitive pressure that schools face from close competitors in their SAT quintile, while overstating demand pressure from cross-groups.

8. Conclusions

We find that students do appear to value several college attributes which we categorize as "consumption" because their benefits arguably accrue only while actually enrolled. For instance, college spending on student activities, sports, and dormitories are significant predictors of college choice, unlike spending on instruction and academic support. Specifically, we estimate that students would be willing to pay 7 percent more to attend a school that spends 100 percent more on student and auxiliary services (dorms, sports, and food service) but is unwilling to spend more to attend a college that spends more on instruction (in fact the point estimate is negative). However, there is significant heterogeneity of preferences across students, with higher achieving students having a greater willingness-to-pay for academic quality than their less academically-oriented peers.

The existence of significant preference heterogeneity has important implications for how colleges can attract students. Our simulations suggest that policies that reallocate financial resources away from these non-academic aspects to instruction would not enable most schools to attract more or better students. However, since student preferences for college attributes are very heterogeneous, different colleges face very different incentives for changing their characteristics depending on their current student body and those they are trying to attract. Less selective schools face higher total enrollment elasticities with respect to spending on services than more selective schools, whereas it is the reverse for spending on instruction and academic quality. In fact, our estimates suggest that less selective schools will actually harm enrollment by improving academic quality. However, in the market for high achieving students, this pattern is much more muted with institutions having comparable incentives for investing in academic quality.

More generally, our results suggest that colleges compete for students on many dimensions – price, consumption amenities, academics – and that different students respond differently to these attributes because preferences are so heterogeneous. This highlights two broad areas for future research. First, it would be natural to extend this analysis to understand the behavior of colleges. In principle one could uncover colleges' objectives, given their actions and the demand-side incentives that student preferences create. For students, our analysis could be extended to understand how differences in preferences influence how students engage with college and persist. Variation in preferences for consumption and academics between students is one possible explanation for differences in college completion that has not been explored.

References

Albouy, David, 2009. “What are Cities Worth? Land Rents, Local Productivity, and the Capitalization of Amenity Values.” NBER WP No. 14981.

Arcidiacono, P. (2004), ”Ability sorting and the returns to college major”, Journal of Econometrics 121, 343-375.

Alstadsæter, A., A.-S. Kolm, and B. Larsen (2008), ”Money or joy: The choice of educational type ,” European Journal of Political Economy 24, 107-122.

Alstadsæter, A., (2009), ”Measuring the Consumption Value of Higher Education.” CESifo Working Paper No. 2799.

Alstadsæter, Annettte, 2003. Income tax, consumption value of education, and the choice of educational type. CESifo Working Paper No. 1055.

Arum, Richard and Josipa Roksa, 2011. *Academically Adrift: Limited Learning on College Campuses.* University of Chicago Press: Chicago, IL.

Avery, Chris, Mark Glickman, Caroline Hoxby, and Andrew Metrick, 2005. A Revealed Preference Ranking of U.S. Colleges and Universities. NBER WP 10803.

Babcock, Philip and Minday Marks. The Falling Time Cost of College: Evidence from Half a Century of Time Use Data , Review of Economics and Statistics, forthcoming

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, 2007. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy,* August 2007.

Becker, Gary (1964). Human Capital. New York: Columbia University Press.

Bound, John, Michael Lovenheim, and Sarah E. Turner. 2007. “Understanding the Decrease in College Completion Rates and the Increased Time to the Baccalaureate Degree.” Population Studies center Research Report 07-626, University of Michigan.

Bowers, T.A., & Pugh, R.C. (1973). Factors underlying college choice by students and parents. Journal of College Student Personnel, 220-224.

Bovenberg, A. Lans, Jacobs, Bas, 2005. Redistribution and educations subsidies are Siamese twins. Journal of Public Economics 89 (11−12), 2005–2035

Carneiro, P., K.T. Hansen, and J.J. Heckman (2003), ”Estimating distributions of treatment effects with and application to the returns to schooling and measurement of the effects of uncertainty of college choice”, International Economic Review 44(2), 361-422.

Chapman, R.C. & Jackson, R. (1987). College choices of academically able students: The influence of no-need financial aid and other factors. Research monograph No. 10. New York: College Board.

Christiansen, C., Joensen, J.S., and Nielsen, H.S. (2007), ”The risk-return trade-off in human capital investment”, Labour Economics 14, 971-986.

Clotfelter, Charles T., 2011. *Big-Time Sports in American Universities*. Cambridge University Press: New York, New York.

Delta Cost Project, 2010. Trends in College Spending 1998-2008: Where Does the Money Come From? Where Does it Go? What Does it Buy? Delta Cost Project: Washington, DC.

Dolton, P.J., G.H. Makepeace, and W. van der Klaauw (1989), ”Occupational choice and earnings determination: The role of sample selection and non-pecuniary factors”, *Oxford Economic Papers* **41**, 573-594.

Donnellan, J. (2002) **–** The impact of marketer controlled factors on college-choice decisions by students at a public research university, Dissertation submitted to the Graduate School of the University of Massachusetts Amherst. UMI number: 3039350.

Duffy, E.A., & Goldberg, I. (1998). College admissions and financial aid, 1955-1994. Princeton, New Jersey: Princeton University Press - notes that by the late 1970s, more colleges were using aggressive marketing strategies to fill their classrooms

Dur, R. and A. Glazer (2008), ”Subsidizing enjoyable education”, *Labour Economics* **15(5)**, 1023-1039.

Daymont, T. and P. Andrisani (1984), ”Job preferences, college major, and the gender gap in earnings”, *Journal of Human Resources* **19**, 408-428.

Drewes and Michael (2006) HOW DO STUDENTS CHOOSE A UNIVERSITY?: An Analysis of Applications to Universities in Ontario, Canada. Research in Higher Education, Vol. 47, No. 7, November 2006.

Epple, Dennis, Richard Romano, and Holger Sieg, 2006. “Admission, Tuition, and Financial Aid Policies in the Market for Higher Education.” *Econometrica* 74 (4): 885-928.

Gale, William G. and John Sabelhaus, 1999. “Perspectives on the Household Saving Rate,” Brookings Papers on Economic Activity, Vol. 1999, No. 1 (1999), pp. 181-224. **–** Discusses the different measures and trends, doesn’t discuss explanations.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, John Sabelhaus, 1996. “Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis.” Brookings Papers on Economic Activity, Vol. 1996, No. 1 (1996), pp. 315-407

Greenspan, Alan and Kennedy, James, 2007, "Sources and Uses of Equity Extracted from Homes." Finance and Economic Discussion Series 2007-20. (Washington: Board of Governors of the Federal Reserve System).

Griliches, Z. "Research Expenditures, Education and the Aggregate Agricultural Production Function." A.E.R. 54 (December 1964): 961-74. .

-- "Notes on the Role of Education in Production Functions and Growth Accounting." In Education, Income, and Human Capital, edited by W. Lee Hansen. New York: Nat. Bur. Econ. Res., 1970.

Gullason, E.T. (1989), The consumption value of schooling. An empirical estimate of one aspect, Journal of Human Resources 24(2), 287-98.

Hastings, Justine S., Thomas J. Kane, and Doug O. Staiger, 2009. "Heterogeneous Preferences and the Efficacy of Public School Choice." Mimeo. May 2009.

Heckman, J.J., L. Lochner, and C. Taber (1999), "General-equilibrium cost-benefit analysis of education and tax policies", In: Ranis, G., L.K. Raut (Eds.): Trade, Growth, and Development. Elsevier, Amsterdam, 291-349.

Heckman, J.J. "A life cycle model of earnings, learning, and consumption." Journal of Political Economy, Vol. 84 (1976).

Hossler, D., Schmit, J. & Vesper, N. (1998). Going to college: Social, economic and educational factors' influence on decisions students make. Baltimore: Johns Hopkins University Press.

Hoxby, Caroline, 1997. How the Changing Market Structure of U.S. Higher Education Explains College Tuition. NBER Working Paper 6323.

Juster, F. Thomas, Joseph P. Lupton, James P. Smith, and Frank Stafford, 2005. "The Decline in Household Saving and the Wealth Effect" Review of Economics and Statistics, Nov 2005 87(4): 20-27.

Keane, Michael P., 2002. Financial aid, borrowing constraints and college attendance: Evidence from structural estimates. American Economic Review 92 (2), 293–297.

Keller, M.J. & McKewon, M.P. (1984). Factors contributing to postsecondary enrollment decisions of Maryland National Merit Scholarship Semifinalists. Paper presented at the annual meeting of the Association for the Study of Higher Education in Chicago

Kirp, David L. (2005). "XXX" In Declining by degrees of mediocrity : higher education at risk / edited by Richard H. Hersch and John Merrow. New York : Palgrave Macmillan, 2005.

Kodde, D.A. and J.M.M. Ritzen (1984), "Integrating consumption and investment motives in a neoclassical model of demand for education", Kyklos 37(4), 598-605.

Lazear, Edward (1977). "Education: consumption or production?" Journal of Political Economy 85(3).

Lin, L. (1997) – "What are student education and educational related needs?", Marketing and Research Today, (August), Vol.25, Nº.3, pp.199-212.

Long, Bridget Terry, 2004. How have college decisions changed over time? An application of the conditional logistic choice model. *Journal of Econometrics* 121 (2004): 271-296.

Malchow-Møller, Nikolaj, Skaksen, Jan Rose, 2004. How to finance education – Taxes or tuition fees? CEBR discussion paper, pp. 2003–2028.

McDuff, DeForest, 2007. Quality, tuition, and applications to in-state public colleges. *Economics of Education Review*. Vol. 26: 433-449.

Michael,R.T.: 'Education in nonmarket production', Journal of Political Economy, Vol. 80 (1973).

Monks, James and Ronald G. Ehrenberg, 1999. "The Impact of *U.S News & World Report* College Rankings On Admission Outcomes and Pricing Decisions at Selective Private Institutions." NBER Working Paper No. 7227 July 1999

Oreopoulos, Philip and Kjell Salvanes, 2011. How large are returns to schooling? Hint: Money isn't everything." *Journal of Economic Perspectives* 25(1): 159-184.

Oosterbeek, H. and H.V. Ophem (2000), "Schooling choices: Preferences, discount rates, and rates of return", Empirical Economics 25, 15-34.

Oosterbeek, H. and D. Webbink (1995), "Enrolment in higher education in the Netherlands", De Economist 143, 367-380.

Parker, Jonathan, 1999. "Spendthrift in America? On Two Decades of Decline in the US Saving Rate." In *NBER Macroeconomics Annual 1999.*

Pope, Devin G. and Jaren C. Pope. 2009. "The Impact of College Sports Success on the Quantity and Quality of Student Applications." Southern Economic Journal (2009) 75(3), 750-780

Pope, Devin G. and Jaren C. Pope. 2008. "Understanding College Application Decisions: Why College Sports Success Matters." mimeo.

Paulo Guimaraes, Pedro Portugal. "A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects." (January 2009)
IZA DP No. 3935

James M. Poterba, 2000. "Stock Market Wealth and Consumption" The Journal of Economic Perspectives, Vol. 14, No. 2 (Spring, 2000), pp. 99-118.

Schaafsma, J. (1976), "The consumption and investment aspects of the demand for education", Journal of Human Resources 11(2), 233-42.

Schultz, T. The Economic Value of Education. New York: Columbia Univ. Press, 1963

Scott-Clayton, Judith (2007), “What Explains Rising Labor Supply Among U.S. Undergraduates, 1970-2003?” Harvard University mimeo.

Stewart, N.R., et al. (1987). Counselor impact on college choice. Paper presented at the annual meeting of the American Educational Research Association in Washington DC.

Train, Kenneth E., 2003. Discrete Choice Methods with Simulation. New York: Cambridge University Press.

Tunis, J.R. (1939). Choosing a college. New York: Harcourt, Brace and Company.

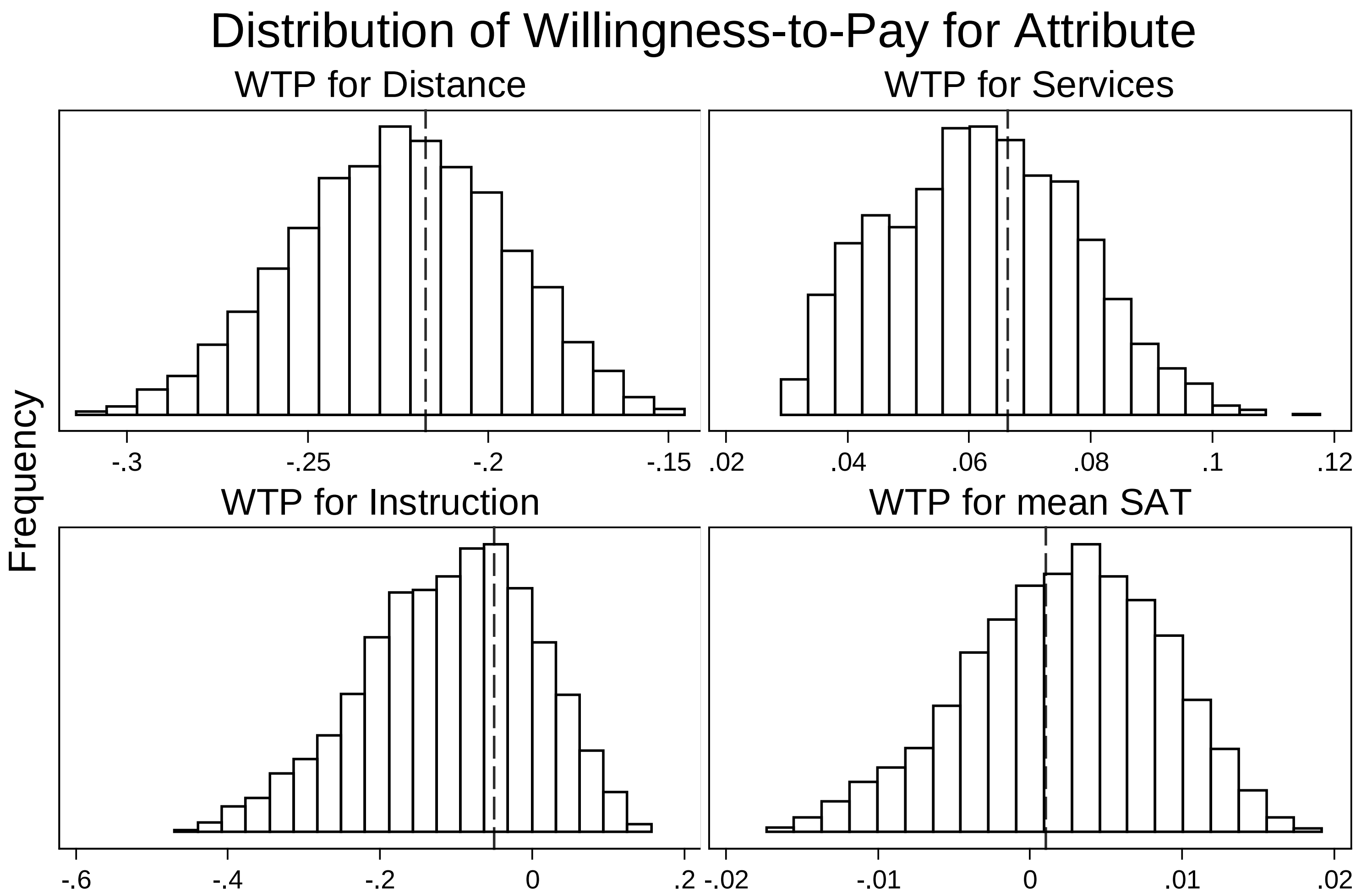
Turner, Sarah (2004), “Going to College and Finishing College: Explaining Different Educational Outcomes,” in College Choices: The Economics of Which College, When College, and How to Pay For It, ed. Caroline Hoxby. Chicago: Chicago University Press.

Weiler, William C. (1996). Factors Influencing the Matriculation Choices of High Ability Students. Economics of Education Review, Vol. 15, No. 1, pp. 23-36, 1996

Welch, F. "Education in Production." J.P.E. 78, no. 1 (January/February 1970): 35-59.

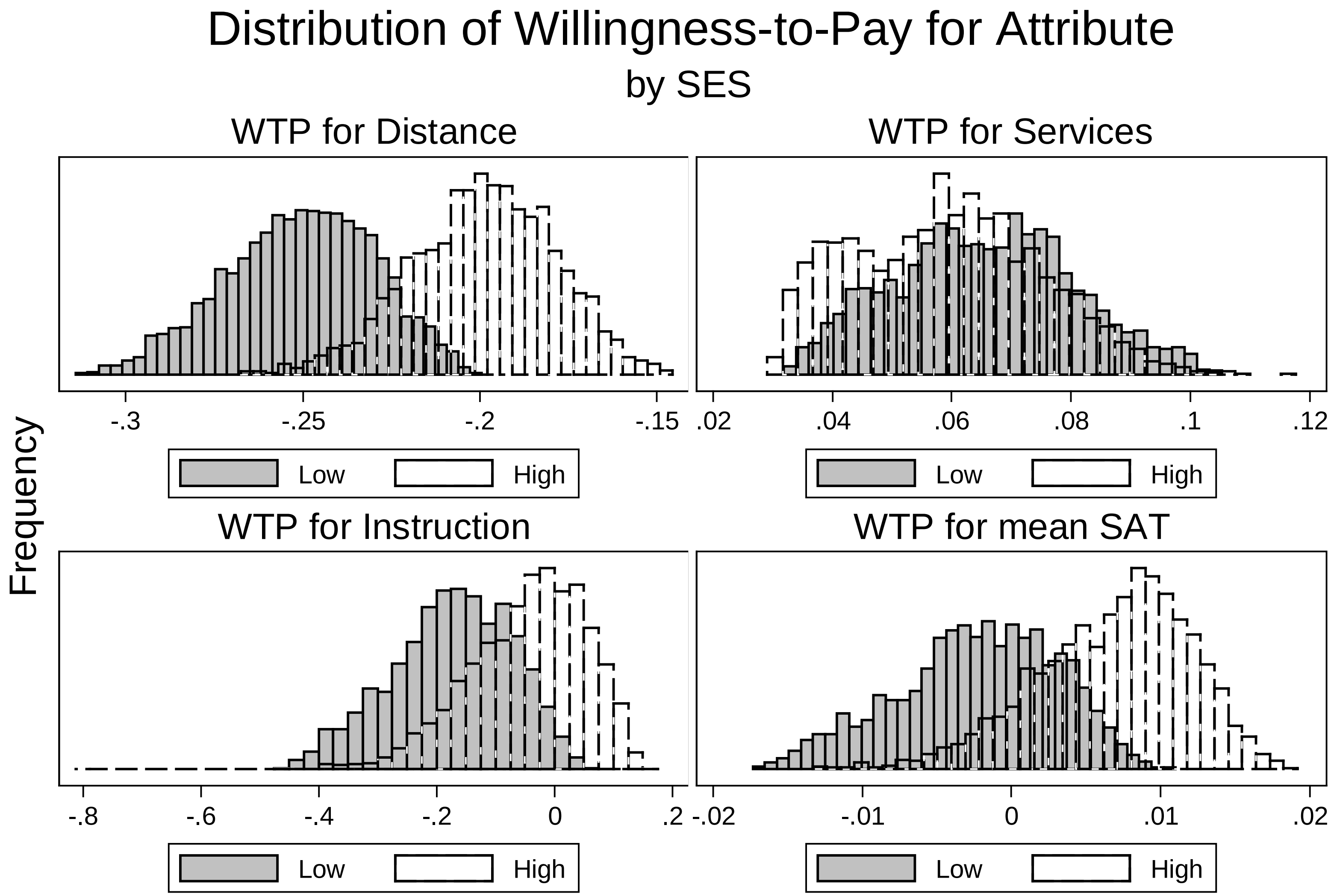
--. "Education, Information and Efficiency." Unpublished Working Paper no. 1, Nat. Bur. Econ. Res., 1973. (b)

Figure 1



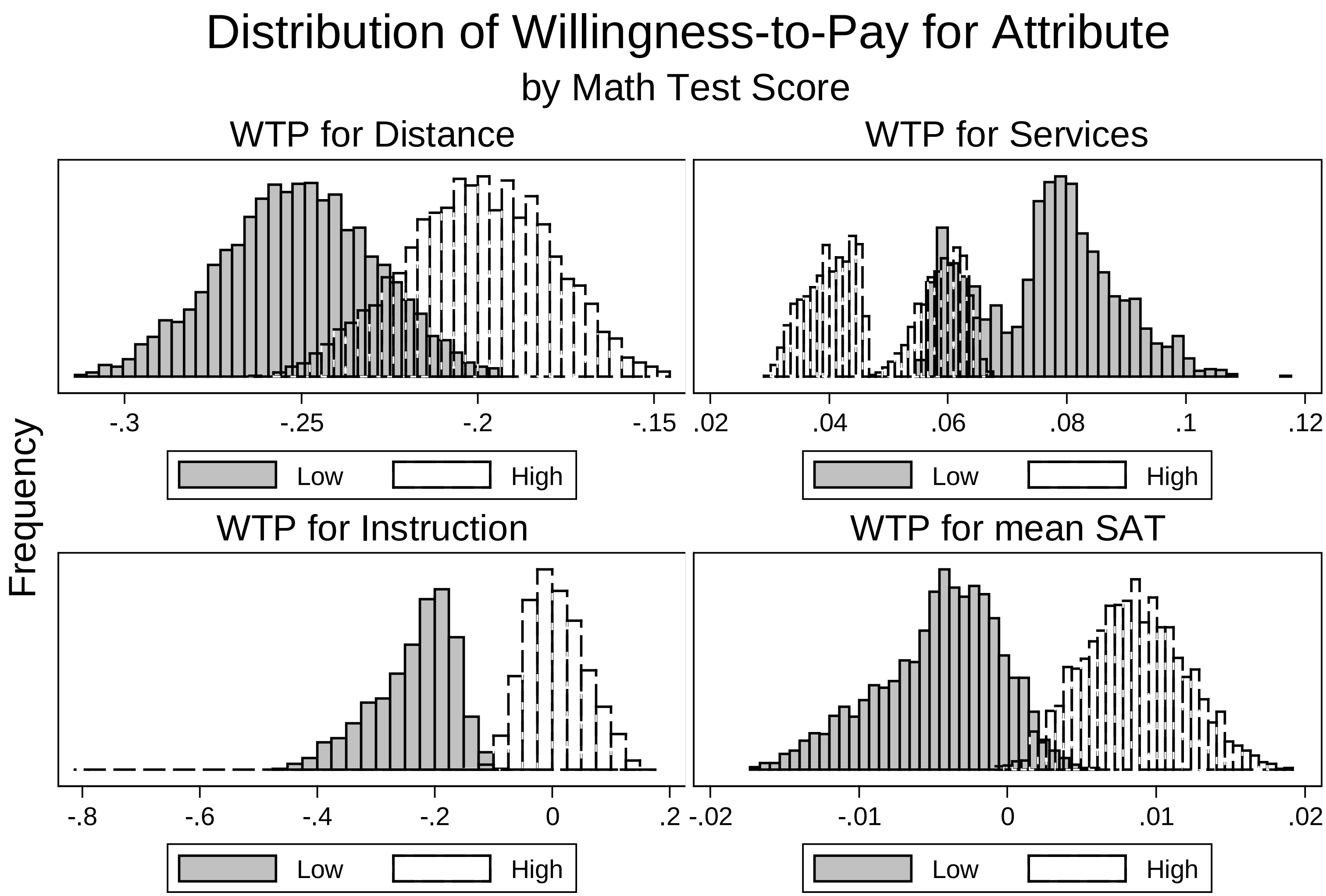
Notes: Estimates come from the model in Table 6 (Specification 2) which includes interactions between college characteristics and male, math score, and SES. Dashed line indicates value for the WTP when heterogeneity is not permitted, estimated in Table 5 (Specification 4).

Figure 2



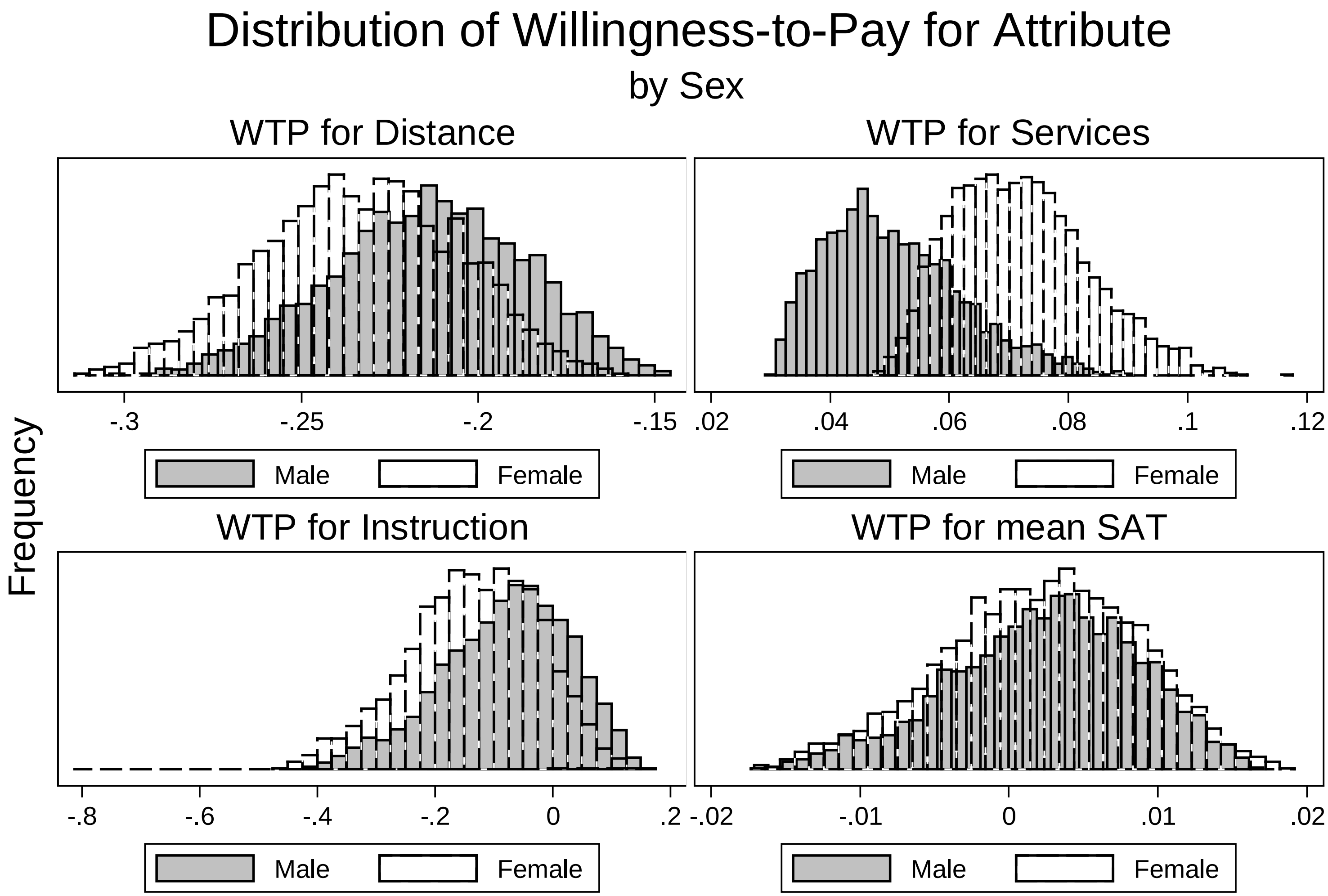
Notes: Estimates come from model with interactions between college characteristics and male, SES, math score
High and low groups represent the top and bottom third by standardized SES index. Middle third omitted.

Figure 3



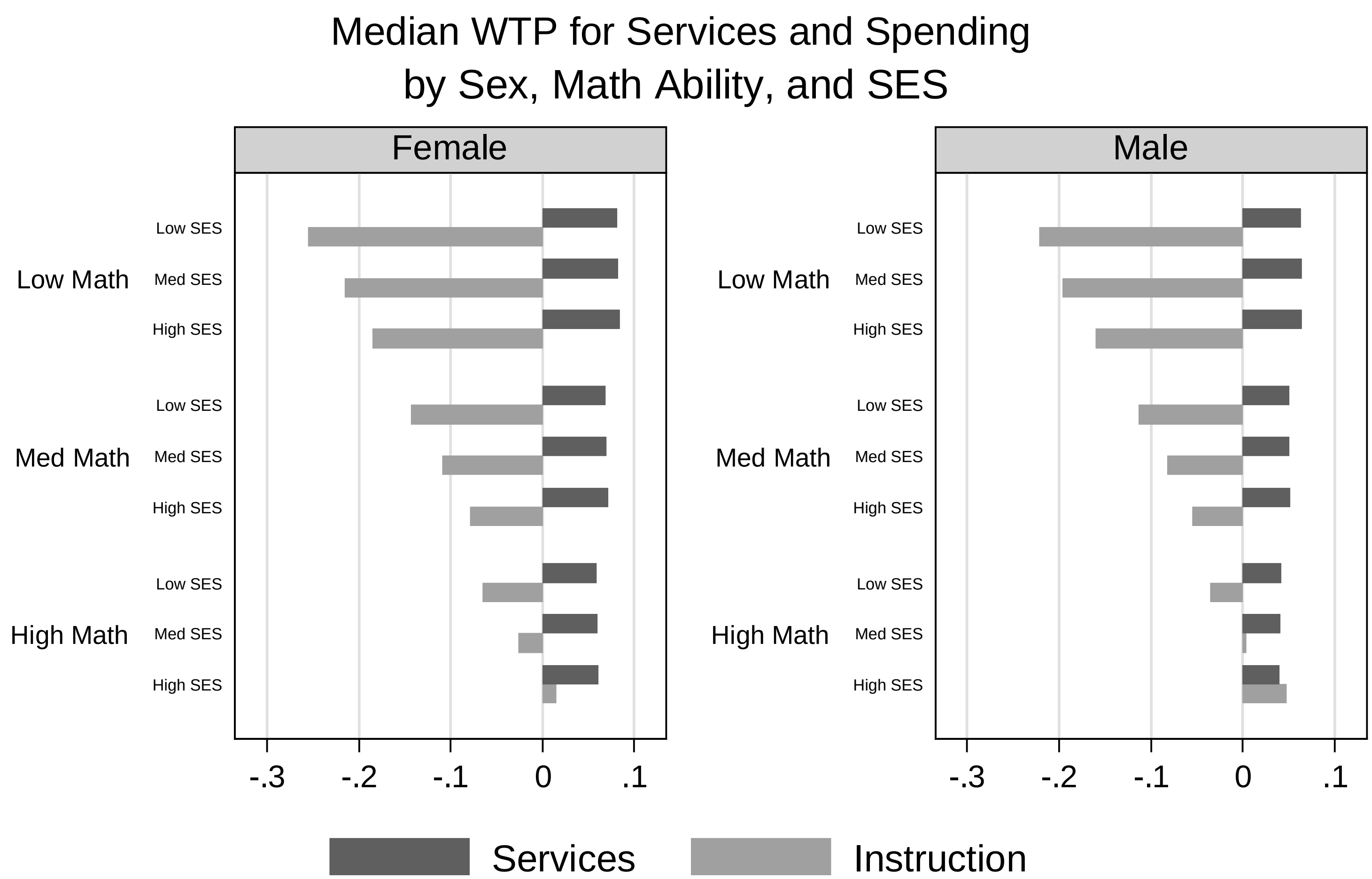
Notes: Estimates come from model with interactions between college characteristics and male, SES, math score
High and low groups represent the top and bottom third by standardized math score. Middle third omitted.

Figure 4



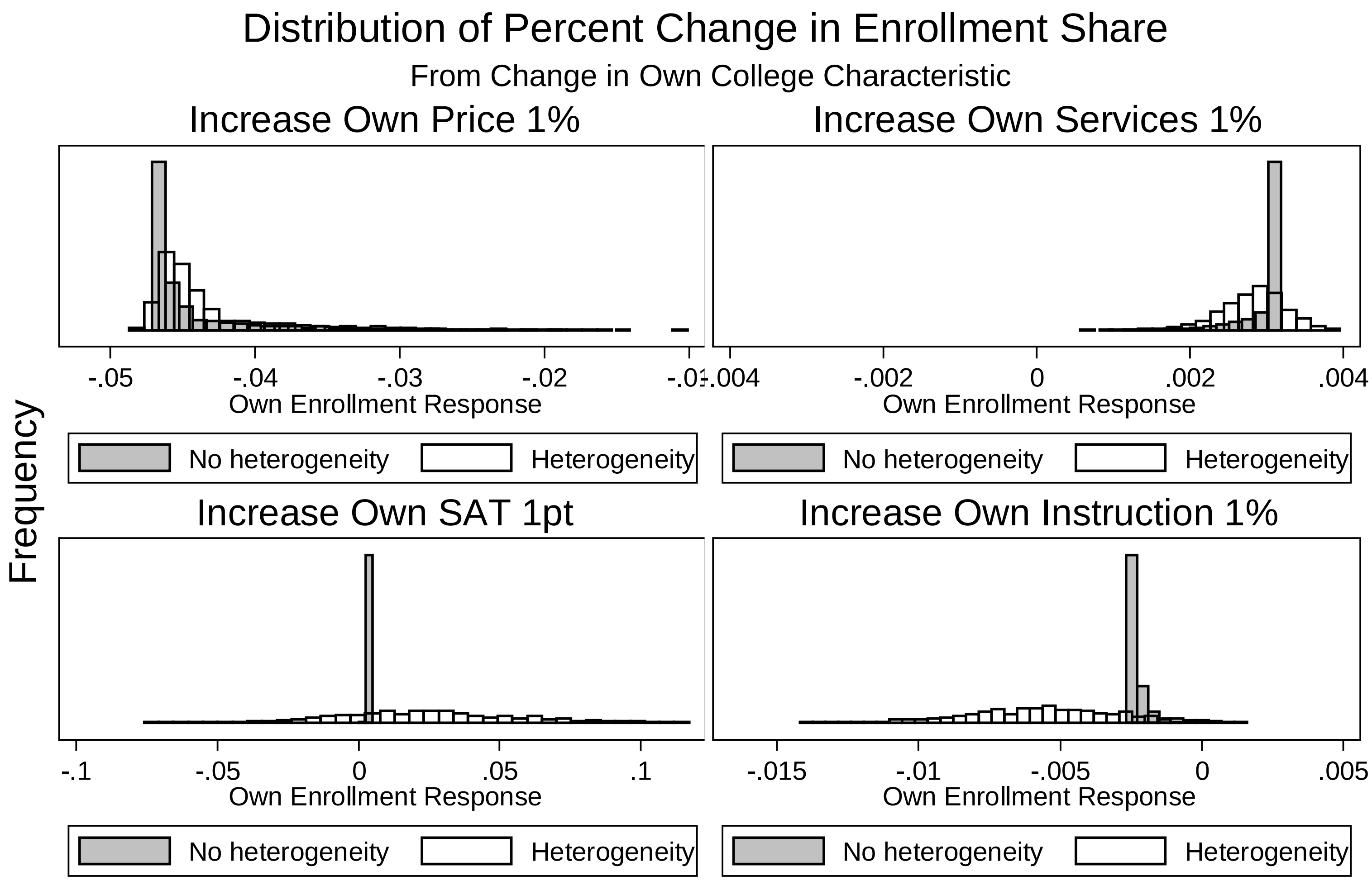
Notes: Estimates come from model with interactions between college characteristics and male, SES, math score
Other notes

Figure 5



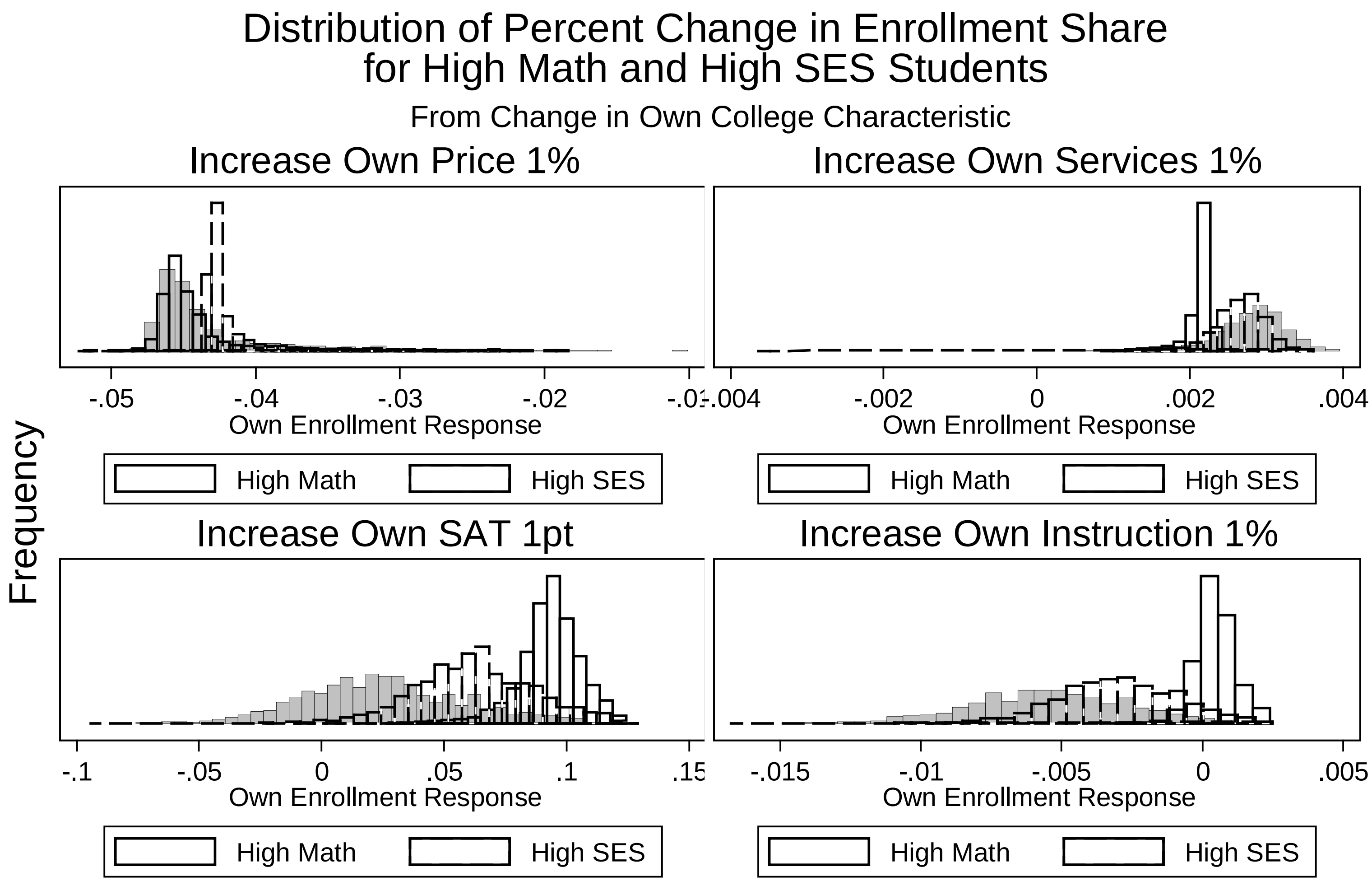
Note: WTP is calculated as minus the ratio of the coefficients on the spending category and cost. Estimates come from the model that includes interactions between the five college characteristics and male, math score, and SES.

Figure 6



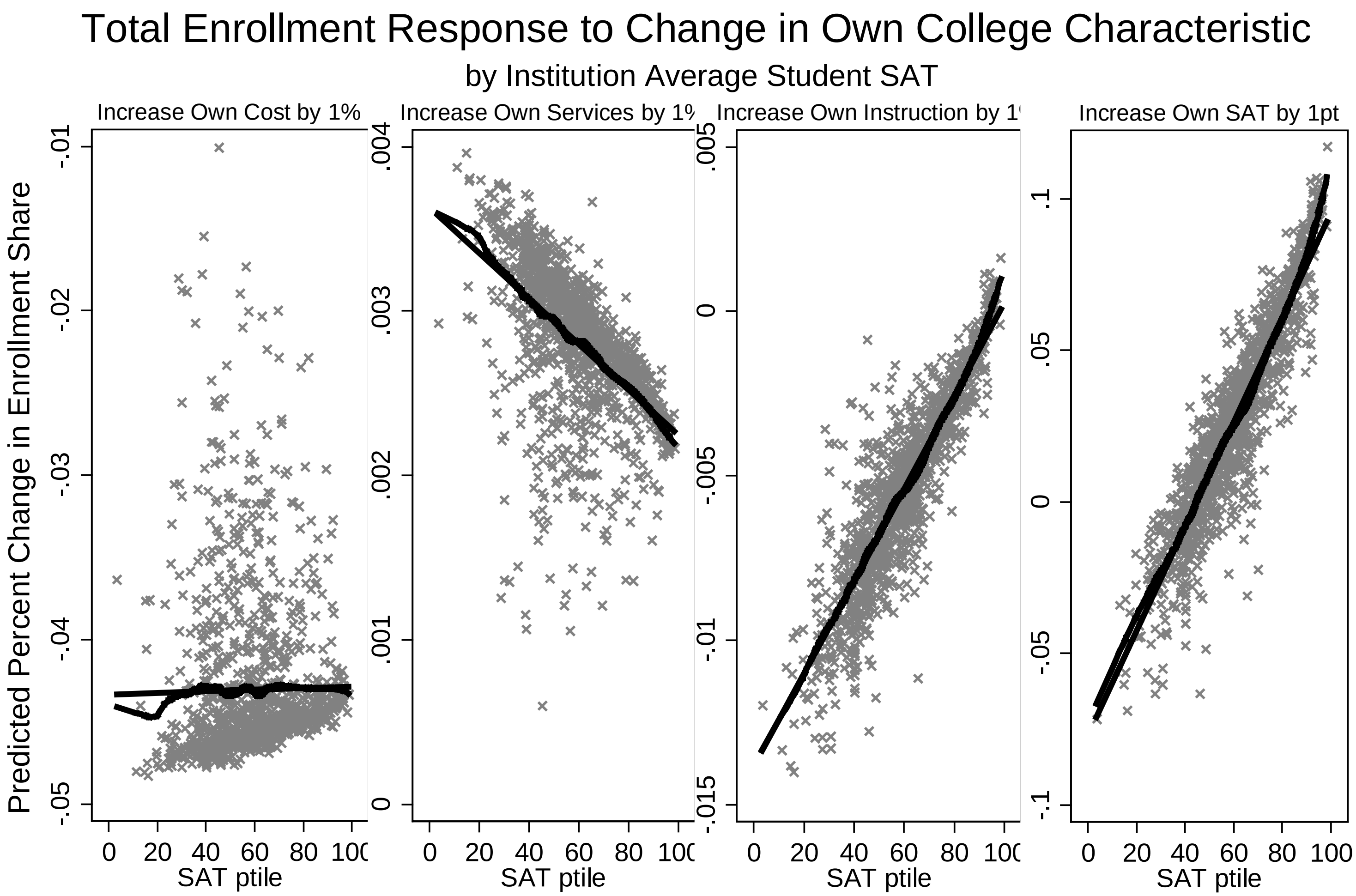
Notes: Each graph plots the distribution of the percent change in enrollment at each individual college if this college were to change a single characteristic.
Enrollment response is simulated using the estimated choice model in Table 6 (Specification 2) which includes interactions between college and student characteristics.

Figure 7



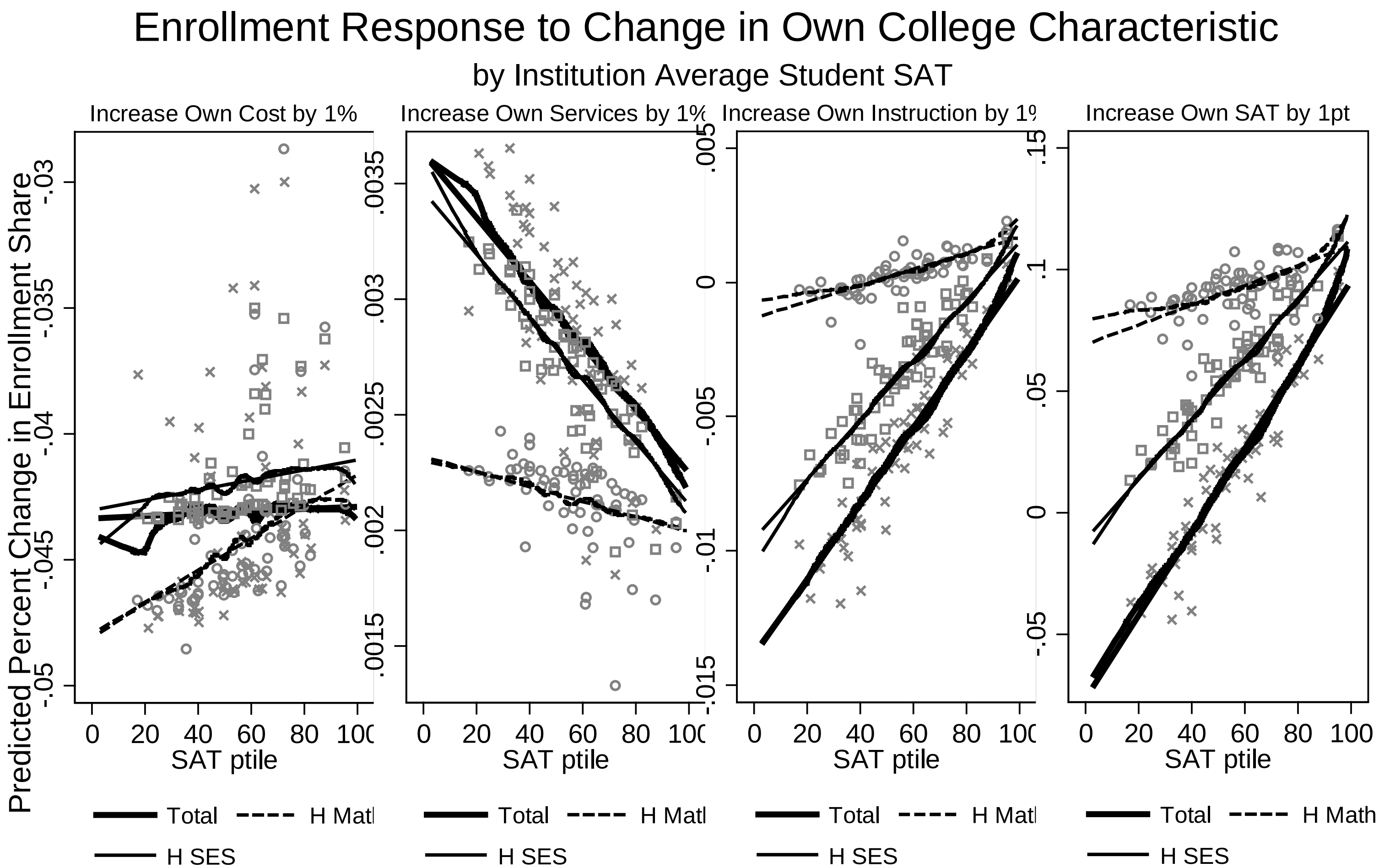
Notes: Each graph plots the distribution of the percent change in enrollment at each individual college if this college were to change a single characteristic.
Enrollment response is simulated using the estimated choice model in Table 6 (Specification 2) which includes interactions between college and student characteristics. Distribution for total enrollment is indicated by shaded distribution in background.

Figure 8



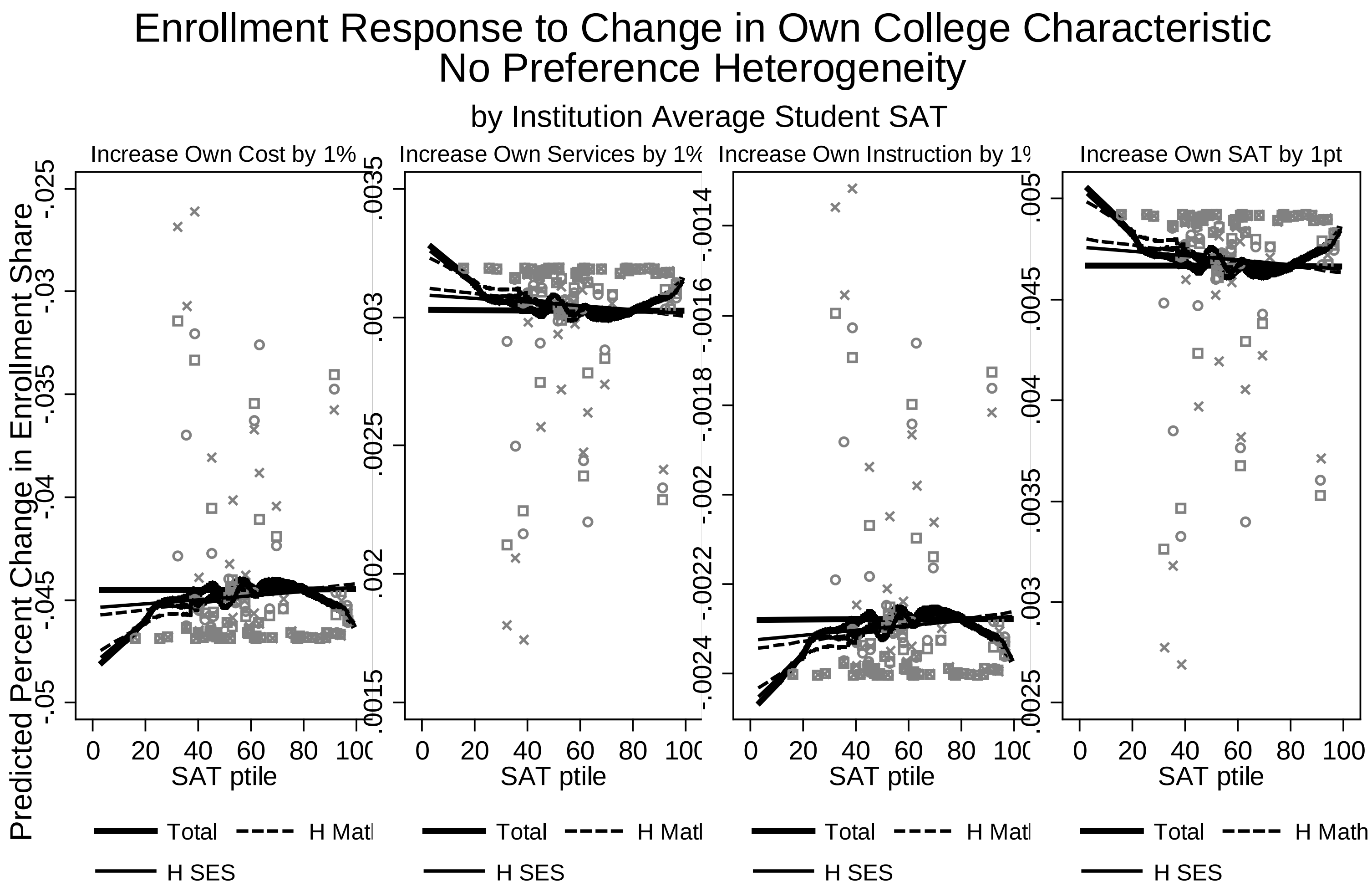
Each point represents a separate simulation where the characteristic of a single college is changed in isolation.
Enrollment response is simulated using the estimated choice model in Table 6 (Specification 2) which includes interactions between college and student characteristics. Graph includes both linear and smoothed prediction lines. Lowess smoothed lines use a bandwidth of 0.20

Figure 9



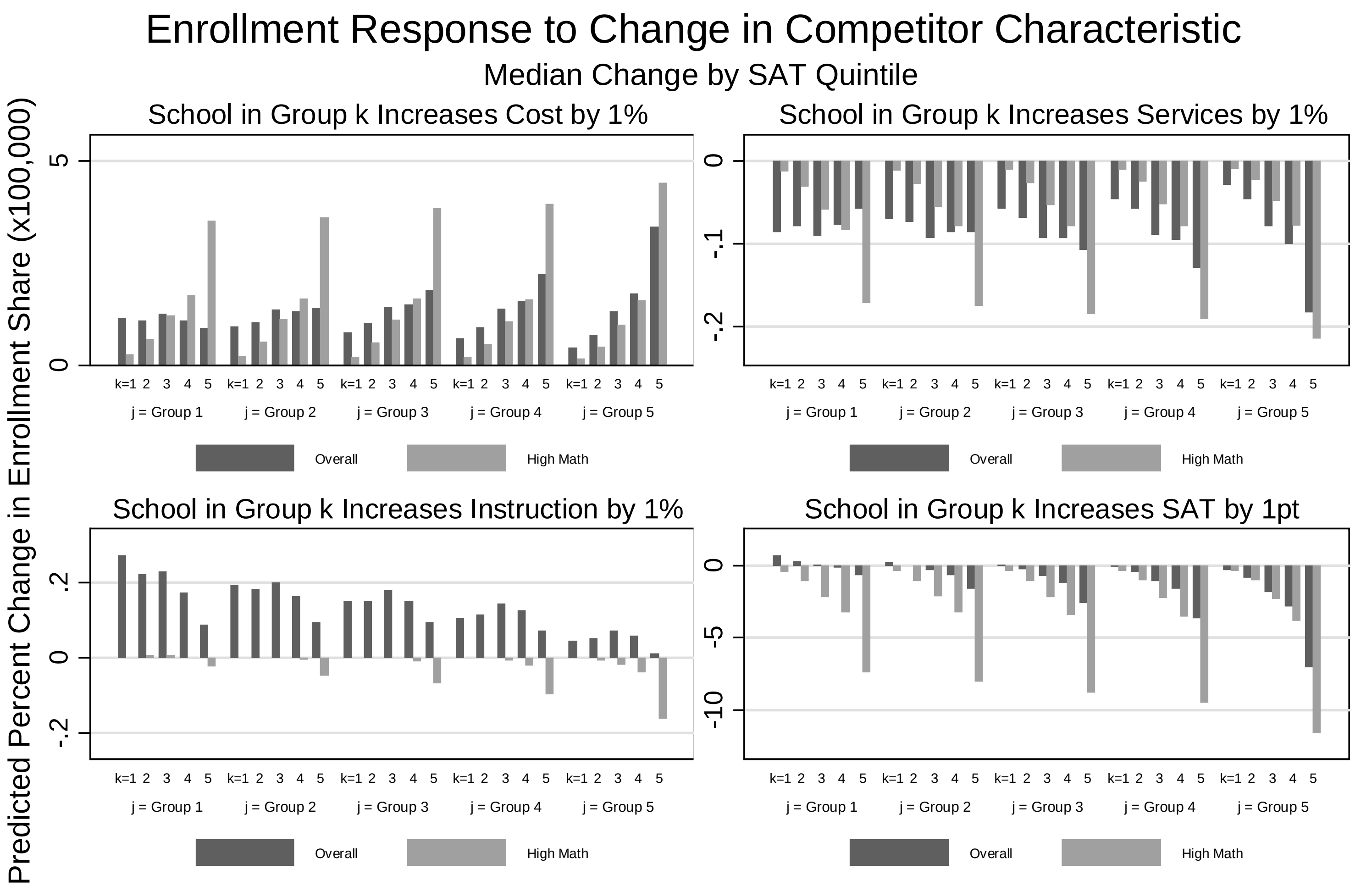
Each point represents a separate simulation where the characteristic of a single college is changed in isolation.
Enrollment response is simulated using the estimated choice model in Table 6 (Specification 2) which includes interactions between college and student characteristics. Graph includes both linear and smoothed prediction lines. Lowess smoothed lines use a bandwidth of 0.20.
For clarity, scatter plot only shows 200 random observations.

Figure 10



Each point represents a separate simulation where the characteristic of a single college is changed in isolation.
Enrollment response is simulated using the estimated choice model in Table 5 (Specification 3).
Lowess smoothed lines use a bandwidth of 0.20. For clarity, scatter plot only shows 200 random observations.

Figure 11



Bars represent the median percent enrollment change at schools in SAT quintile j when a single school in SAT quintile k (along x-axis) changes a characteristic. Median is calculated across all pairs in the j-k cell. Sample includes all 1261 X 1260 cross elasticitie Enrollment response is simulated using the estimated choice model in Table 6 (Specification 2) which includes interactions between college and student characteristics.

Table 1. College Summary Statistics

	1992		2004	
mean (std. dev.)	public	private	public	private
In-State Tuition	3,584	14,712	5,560	20,827
	(1,476)	(5,647)	(2,356)	(7,302)
Out-of-State Tuition	9,191	14,773	13,667	20,869
	(3,177)	(5,575)	(4,504)	(7,246)
Room and Board	5,349	6,021	7,098	7,323
	(1,643)	(1,844)	(1,504)	(1,776)
Freshmen Fall Enrollment	1,307	429	1,607	501
	(1,051)	(448)	(1,436)	(504)
Full-Time Equivalent Enrollment	8,513	2,251	9,649	2,853
	(7,695)	(2,950)	(8,989)	(3,552)
Instructional and Academic Support $ per FTE	7,751	8,049	8,490	10,075
	(3,574)	(4,548)	(3,647)	(5,359)
Student Services and Auxiliary Support $ per FTE	3,394	5,236	3,726	6,438
	(1,565)	(2,624)	(1,794)	(2,880)
Median or mean SAT Ptile	57.97	64.47	51.77	58.79
	(15.99)	(17.40)	(16.06)	(18.63)
Highest degree offered is BA	0.17	0.41	0.17	0.31
	(0.38)	(0.49)	(0.37)	(0.46)
Highest degree offered is MA	0.47	0.43	0.42	0.46
	(0.50)	(0.50)	(0.49)	(0.50)
Highest degree offered is PhD	0.36	0.16	0.41	0.23
	(0.48)	(0.37)	(0.49)	(0.42)
Number of Schools	530	879	570	887

Notes: All spending variables are deflated by the CPI-U and are in 2009 dollars.

Table 2. Pair-wise Correlations of College Characteristics

	Log In-State Tuition + RBR	Log Out-of-State Tuition + RBR	Log Services Spending	Log Instructional Spending	Mean SAT
	Correlations in 2004				
Log In-State Tuition + RBR	1.000				
Log Out-of-State Tuition + RBR	0.861	1.000			
Log Services Spending	0.633	0.588	1.000		
Log Instructional Spending	0.513	0.648	0.597	1.000	
Mean SAT	0.492	0.608	0.489	0.634	1.000
	Correlations in 1992				
Log In-State Tuition + RBR	1.000				
Log Out-of-State Tuition + RBR	0.889	1.000			
Log Services Spending	0.580	0.592	1.000		
Log Instructional Spending	0.434	0.602	0.547	1.000	
Mean SAT	0.483	0.564	0.491	0.588	1.000
	Correlation of difference 2004-1992				
Log In-State Tuition + RBR	1.000				
Log Out-of-State Tuition + RBR	0.845	1.000			
Log Services Spending	0.111	0.086	1.000		
Log Instructional Spending	0.091	0.076	0.457	1.000	
Mean SAT	*0.019*	*0.019*	*0.019*	*0.067*	1.000

Notes: Each cell is the college-level pair-wise (unweighted) correlation between each pair of variables. Correlations where observations are weighted based on the number of individuals choosing the school in our sample are very similar, both qualitatively and quantitatively. Estimates in italics indicate correlation is not significant at the 95% level. All other correlations are significant.

Table 3: Student Characteristics

	1992		2004	
Number of students in our analysis sample	4,088		5,753	
Background Characteristics of Analysis Sample	Mean	SD	Mean	SD
Male	0.46	0.50	0.45	0.50
Standardized math score	0.62	0.83	0.65	0.82
Standardized SES	0.41	0.97	0.48	0.97
*Standardized composite measure of importance of various college characteristics in analysis sample**				
Academics (courses, reputation)	0.27	0.74	0.33	0.69
Cost (low costs, availability of financial aid)	-0.14	0.65	-0.02	0.67
Social Life (athletics, social life)	-0.03	0.83	0.18	0.87
Characteristics of institution student attended				
Cost (Tuition + Fees + Room and Board)	14,801	8,608	20,859	10,577
Distance from institution to home (miles)	196	389	219	481
School Mean SAT (percentile)	67.57	17.28	62.08	17.14
Spending on instruction/fte ($2009)	9,990	6,836	11,855	9,061
Spending on student services/fte ($2009)	4,646	2,630	5,286	3,438
Log(enrollment)	7.10	0.97	7.34	0.95
Predicted probability of admission	0.71	0.15	0.81	0.18
Predicted net price	11,404	6,573	14,892	7,430
In state	0.74	0.44	0.73	0.44
In region	0.82	0.39	0.82	0.38
Characteristics of institutions not attended				
Cost (Tuition + Fees + Room and Board)	18,694	6,948	26,014	8,396
Distance from institution to home (miles)	954	709	996	778
School Mean SAT (percentile)	64.48	17.13	57.84	18.06
Spending on instruction/fte ($2009)	8,644	5,643	10,642	8,396
Spending on student services/fte ($2009)	4,685	2,598	5,678	3,538
Log(enrollment)	6.45	0.93	6.57	0.94
Predicted probability of admission	0.70	0.18	0.81	0.22
Predicted net price	13,288	5,308	17,003	6,509
In state	0.03	0.18	0.04	0.18
In region	0.12	0.33	0.13	0.33

*Simple item average, standardized with 1972 mean

Table 4: Conditional Logit Estimates of the Predictors of College Choice, Separate Cross-sections

	High School Graduates in 1992				High School Graduates in 2004			
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log (Tuition, Fees, Room & Board)	0.164***	0.132***	0.147***	0.139***	0.162***	0.148***	0.166***	0.146***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
Log (Distance)	0.322***	0.318***	0.316***	0.320***	0.340***	0.337***	0.336***	0.339***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Log (Spending on Student Services/fte)		2.452***	2.673***	2.508***		1.611***	1.819***	1.607***
		(0.118)	(0.128)	(0.123)		(0.053)	(0.058)	(0.053)
Log (Spending on Instruction/fte)	1.987***	1.295***	1.471***	1.453***	1.506***	1.192***	1.444***	1.559***
	(0.088)	(0.068)	(0.074)	(0.079)	(0.055)	(0.049)	(0.056)	(0.069)
School Mean SAT (percentile)	1.017***	1.012***		1.018***	1.019***	1.015***		1.023***
	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)		(0.001)
Log (Lagged first time freshman enrollment)	1.659***	1.819***	1.899***	1.787***	1.893***	2.036***	2.200***	2.046***
	(0.037)	(0.042)	(0.043)	(0.040)	(0.032)	(0.037)	(0.040)	(0.036)
Institution state unemployment rate	0.965***	0.978*	0.970**	0.989	1.004	0.989	0.967*	0.970
	(0.012)	(0.013)	(0.012)	(0.013)	(0.019)	(0.019)	(0.018)	(0.019)
Accounting for probability of admissions	No	No	No	Yes	No	No	No	Yes
Number of observations	3,989,091	3,989,091	3,989,091	3,989,087	6,361,028	6,361,028	6,361,028	6,361,028

*** p<0.01, ** p<0.05, * p<0.1

Notes: Odds ratios and robust standard errors shown above.

Table 5: Conditional Logit Estimates of the Predictors of College Choice, No Preference Heterogeneity (Odds Ratios Reported)

	Dept Variable: College Chosen by High School Graduates in 1992 and 2004						
Independent Variables	(1)	(1)	(2)	(3)	(4)	(5)	(6)
Log (Tuition, Fees, Room & Board)	0.106 ***	0.143 ***	0.145 ***	0.008 ***	0.008 ***	0.083 ***	0.053 ***
	(0.0025)	(0.0041)	(0.0042)	(0.0006)	(0.0006)	(0.0076)	(0.0034)
Log (Distance)	0.331 ***	0.329 ***	0.331 ***	0.350 ***	0.351 ***	0.468 ***	0.309 ***
	(0.0019)	(0.0020)	(0.0020)	(0.0033)	(0.0033)	(0.0050)	(0.0027)
Log (Spending on Student Services/fte)	1.229 ***	1.854 ***	1.865 ***	1.378 ***	1.367 ***	1.238 **	1.236 **
	(0.0269)	(0.0499)	(0.0509)	(0.1316)	(0.1306)	(0.1083)	(0.1133)
Log (Spending on Instruction/fte)	1.753 ***	1.223 ***	1.512 ***	0.785 *	0.786 *	0.815	0.760 **
	(0.0476)	(0.0392)	(0.0515)	(0.1130)	(0.1129)	(0.1078)	(0.1064)
School Mean SAT (percentile)	1.025 ***	1.014 ***	1.021 ***	1.005 *	1.005	1.005 **	1.016 ***
	(0.0008)	(0.0009)	(0.0009)	(0.0030)	(0.0030)	(0.0026)	(0.0028)
Institution state unemployment rate					1.038	0.978	
					(0.0261)	(0.0267)	
Log(high school grads in institution state)					1.212	1.136	
					(0.3333)	(0.3428)	
College located in the student's home state						4.799 ***	
						(0.2513)	
College located in the student's census region						1.988 ***	
						(0.0900)	
Log (Lagged first time freshman enrollment)	No	Yes	Yes	No	No	No	No
Accounting for Probability of Admissions	No	No	Yes	Yes	Yes	Yes	Yes
College Fixed Effects	No	No	No	Yes	Yes	Yes	Yes
Log (Predicted net price) used as cost measure	No	No	No	No	No	No	Yes
Number of observations	10,350,115	10,350,115	10,350,115	10,350,115	10,350,115	10,350,115	10,350,115
Willingness-to-Pay (s.e.)							
Log (Distance)	-0.491	-0.572	-0.573	-0.217	-0.217	-0.305	-0.401
	(0.0065)	(0.1010)	(0.0101)	(0.0041)	(0.0041)	(0.0107)	(0.0089)
Log (Spending on Student Services/fte)	0.092	0.318	0.323	0.066	0.065	0.086	0.072
	(0.0108)	(0.0144)	(0.0146)	(0.0179)	(0.0179)	(0.0338)	(0.0288)
Log (Spending on Instruction/fte)	0.250	0.104	0.214	-0.050	-0.050	-0.083	-0.094
	(0.0130)	(0.0163)	(0.0169)	(0.0271)	(0.0270)	(0.0512)	(0.0440)
School Mean SAT (percentile)	0.011	0.007	0.011	0.001	0.001	0.002	0.005
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0005)	(0.0010)	(0.0009)

*** p<0.01, ** p<0.05, * p<0.1

Notes: Odds ratios and robust s.e. shown above. Spending on student services also includes spending on auxilary enterprises (primarily food service and dorms). Instruction includes both instruction and academic support services.

Table 6: Conditional Logit Estimates of the Predictors of College Choice, Heterogeneity by Observable Student Characteristic

Independent Variables	Dept Variable: College Chosen by High School Graduates in 1992 and 2004								
	(1)			(2)			(3)		
	Est.		(S.E.)	Est.		(S.E.)	Est.		(S.E.)
Log (Tuition, Fees, Room & Board)	-2.076	***	(0.0505)	-4.812	***	(0.0875)	-3.987	***	(0.0872)
X male	-0.206	***	(0.0635)	-0.217	***	(0.0708)	-0.183	**	(0.0821)
X math score (standardized)	0.058		(0.0437)	0.011		(0.0517)	-0.038		(0.0611)
X SES (standardized)	0.337	***	(0.0368)	0.301	***	(0.0415)	0.326	***	(0.0493)
Log (Distance)	-1.248	***	(0.0106)	-1.216	***	(0.0132)	-1.268	***	(0.0128)
X male	0.015		(0.0125)	0.014		(0.0136)	0.008		(0.0135)
X math score (standardized)	0.080	***	(0.0086)	0.098	***	(0.0095)	0.096	***	(0.0095)
X SES (standardized)	0.144	***	(0.0074)	0.153	***	(0.0081)	0.158	***	(0.0079)
Log (Spending on Student Services/fte)	0.634	***	(0.0401)	0.380	***	(0.1070)	0.211	**	(0.1058)
X male	-0.069		(0.0514)	-0.074		(0.0619)	-0.115	*	(0.0588)
X math score (standardized)	-0.006		(0.0332)	-0.069		(0.0444)	-0.158	***	(0.0420)
X SES (standardized)	0.041		(0.0293)	-0.011		(0.0363)	0.289	***	(0.0343)
Log (Spending on Instruction/fte)	-0.277	***	(0.0534)	-1.019	***	(0.1573)	-1.078	***	(0.1583)
X male	0.104	*	(0.0625)	0.095		(0.0700)	0.096		(0.0716)
X math score (standardized)	0.491	***	(0.0418)	0.581	***	(0.0539)	0.664	***	(0.0548)
X SES (standardized)	0.109	***	(0.0365)	0.152	***	(0.0428)	0.031		(0.0436)
School Mean SAT (percentile)	0.007	***	(0.0013)	-0.012	***	(0.0033)	-0.006	*	(0.0033)
X male	-0.005	***	(0.0018)	-0.006	***	(0.0020)	-0.006	***	(0.0021)
X math score (standardized)	0.025	***	(0.0012)	0.031	***	(0.0015)	0.040	***	(0.0015)
X SES (standardized)	0.009	***	(0.0010)	0.011	***	(0.0012)	0.014	***	(0.0012)
Log (Lagged first time freshman enrollment)	Yes			No			No		
Accounting for Probability of Admissions	Yes			Yes			Yes		
College Fixed Effects	No			Yes			Yes		
Log (Predicted net price) used as cost measure	No			No			Yes		
Number of observations	10,350,115			10,350,115			10,350,115		

*** p<0.01, ** p<0.05, * p<0.1

Notes: Odds ratios and robust s.e. shown above. Spending on student services also includes spending on auxilary enterprises (primarily food service and dorms). I includes both instruction and academic support services.

Table 7: Conditional Logit Estimates of the Predictors of College Choice, Heterogeneity by Stated Preference

11

	Dept Variable: College Chosen by High School Graduates in 1992 and 2004		
Independent Variables	(1)		
	Est.		(S.E.)
Log (Tuition, Fees, Room & Board)	-4.918	***	(0.0813)
X social life important (standardized)	0.033		(0.0253)
X expenses important (standardized)	-0.256	***	(0.0345)
X academics important (standardized)	0.112	***	(0.0309)
Log (Distance)	-1.101	***	(0.0105)
X social life important (standardized)	0.092	***	(0.0080)
X expenses important (standardized)	-0.257	***	(0.0110)
X academics important (standardized)	0.039	***	(0.0103)
Log (Spending on Student Services/fte)	0.278	***	(0.0991)
X social life important (standardized)	0.122	***	(0.0356)
X expenses important (standardized)	-0.147	***	(0.0504)
X academics important (standardized)	0.031		(0.0446)
Log (Spending on Instruction/fte)	-0.338	**	(0.1485)
X social life important (standardized)	-0.097	**	(0.0398)
X expenses important (standardized)	-0.305	***	(0.0563)
X academics important (standardized)	0.275	***	(0.0512)
School Mean SAT (percentile)	0.003		(0.0031)
X social life important (standardized)	0.000		(0.0011)
X expenses important (standardized)	-0.019	***	(0.0015)
X academics important (standardized)	0.005	***	(0.0013)
Accounting for Probability of Admissions	Yes		
College Fixed Effects	Yes		
Log (Predicted net price) used as cost measure	No		
Number of observations	10,350,115		

*** p<0.01, ** p<0.05, * p<0.1

Notes: Odds ratios and robust s.e. shown above. Spending on student services also includes spending on auxilary enterprises (primarily food service and dorms). Instruction includes both instruction and academic support services. Stated preference is constructed by combining answers to several questions about the importance of various factors in college decision into three categories: social life (including athletics), costs (low cost, availability of financial aid), and academics (course offerings and reputation).

Table 8. Own and Cross Enrollment Response to Change in College Characteristic

	Percent change in enrollment share (0.01 = 1% change in enrollment share)							
	Own Response (n= 1261)				Cross Response (n = 1,588,860)			
	Mean	StdDev	Min	Max	Mean	StdDev	Min	Max
Price Increase 1%								
Total	-0.043	0.005	-0.048	-0.010	0.0000342	0.0001584	0	0.0201
High math	-0.044	0.004	-0.052	-0.018	0.0000351	0.0001596	0	0.0306
High SES	-0.042	0.003	-0.044	-0.022	0.0000333	0.0001118	0	0.0129
Services Increase 1%								
Total	0.003	0.0004	0.001	0.004	-0.0000022	0.0000107	-0.0017	0
High math	0.002	0.0002	0.001	0.003	-0.0000017	0.0000076	-0.0018	0
High SES	0.003	0.0003	0.001	0.004	-0.0000021	0.0000072	-0.0013	0
Instruction Increase 1%								
Total	-0.005	0.003	-0.014	0.002	0.0000044	0.0000282	-0.0002	0.0057
High math	0	0.001	-0.003	0.002	-0.0000004	0.0000041	-0.0005	0.0022
High SES	-0.003	0.002	-0.011	0.002	0.0000024	0.000012	-0.0003	0.0045
SAT Increase 1 point								
Total	0.025	0.033	-0.072	0.117	-0.0000198	0.0001495	-0.0189	0.0267
High math	0.093	0.013	0.03	0.124	-0.0000736	0.0003011	-0.0374	0
High SES	0.062	0.024	-0.024	0.129	-0.0000493	0.0001835	-0.0231	0.0167

Notes: Enrollment response is simulated using the estimated choice model which includes interactions between college and student characteristics.

Appendix A: Data and Sample

The student-level data for this analysis is drawn from two datasets collected by the U.S. Department of Education: the National Educational Longitudinal Survey (NELS), which tracks the high school graduating class of 1988 and the Educational Longitudinal Survey (ELS), which tracks the high school graduating class of 2004. Both datasets provided detailed information on student demographics, prior achievement, college application and admission decisions and college enrollment.

Construction of Our Analysis Sample

We include in our analysis only students who we observe enrolled in an "eligible" four-year, not-for-profit college within two years of expected high school graduation. As discussed in the paper, we limit our college sample in several ways to facilitate our focus on amenities arguably related to direct, immediate consumption value. First, we limit our sample to public and non-profit private undergraduate four-year schools only, excluding all two-year (or less) schools, all for-profit schools, and schools offering professional degrees only. Second, we drop specialized divinity, law, medical, specialized health (e.g. nursing), and art schools, though we keep engineering, teaching, military, and business schools. We drop schools with an average of fewer than 50 freshmen or 300 FTEs over our four sample years in an effort to eliminate remaining specialized schools which are arguably not in many students' choice set. We drop from our analysis any school for which we do not have information on instructional spending, student service spending, tuition or room and board costs, zip code, enrollment, or average SAT score. Finally, because they will not contribute at all to the estimation, colleges that were not attended by at least one student in our micro-data sample are dropped (Table A1).

Our data on enrollment school comes from student surveys administered in 1994 for the NELS cohort and in 2006 for the ELS cohort. We define a student's choice school as the first institution she or he attended, according to NELS and ELS surveys. For NELS, students were asked which schools they attended in a 1994 follow-up survey. This is separate from the application survey questions in 1992 asking students in their senior year of high school which post-secondary institutions they applied to and whether they were accepted. The ELS asked students in 2006, two years after graduation, to which schools they applied, were accepted, and attended.

Using the enrollment dates provided in the data, we identify the first institution each student attended. Note that it is possible that we dropped students who began their post-secondary education at an ineligible school, but transferred to an eligible school – even as early as the first Fall following the student's senior year in high school. We plan to change this in future versions of the paper. In the NELS (ELS), this was determined by the IPEDS code listed in *unitid1* (*f2iiped1* if *f2iattnd1*=1).

Our student sample begins with all of the students in the nationally representative set of 12th graders in 1992 (NELS) and 2004 (ELS). Note that the NELS (ELS) starts by surveying students in 8th (10th) grade, but "freshen" their sample to obtain a nationally representative set of

12^{th} graders in the years above. We first drop students who did not first attend one of the eligible institutions in our sample. We then drop students who have missing information on high school state, socioeconomic status, standardized math score, gender, or race. Next, we drop schools from the students' choice set which have missing covariates such as instructional and student spending, tuition and room-and-board costs, enrollment, or average SAT score. Table A2 shows how the sample size changes for each step in the process above.

Finally, we drop from our analysis any student whose choice school was subsequently dropped due to those aforementioned missing covariates. It is possible that a student chose/attended an otherwise-eligible school which was missing a key covariate, such as mean SAT or tuition costs. When this student's choice school was dropped for missing these variables, we dropped the student entirely from the analysis set.

Variable Construction

Finally, for some of our analyses, we also use data on the quality of life and cost-of-living in the geographic areas in which each college is located. Quality of life is measured both at the county level and the consistent Public Use Micro Area (PUMA) level, and is calculated using data from the 2000 census (Albouy 2009). In essence, these hedonic measures incorporate information on local land values, wage levels and housing costs. For this analysis, we use the consistent PUMA quality of life. The cost-of-living index is based on the cost of a weighted bundle of consumer goods. The data is collected annually for a variety of cities across the United States by the Council for Community and Economic Research and local Chambers of Commerce.

From IPEDS, we have a single zip code associated with each institution. For the most part, the institution occupies a space inside the zip code. However, there are also "unique" zip codes, which the US Post Office assigns to institutions (UCLA, for example) which receive large amounts of mail. In these cases, the zip code is associated with a point, often an administrative building or campus post office (http://www.census.gov/geo/ZCTA/zctafaq.html#Q10). We then utilized the Missouri Data Center's Dexter Database to link zip codes to Public Use Micro Areas (PUMA), which were then aggregated into Consistent Public Use Micro Areas (CPUMA). The QOL measure we used was an aggregate for the CPUMA.

To get a measure of urban area for the institution, we utilize ArcGis and Census Tiger/Line files, mapping the coordinates associated with the population-weighted center of the institution's zip code to the closest urban area or cluster, and micro- or metropolitan area. Finally, we assign an institution as urban if the zip code center falls into an urban area or cluster. Approximately 94% of schools were located in a metropolitan or micropolitan area. As noted above, for the 17% of institutions that were not located in an urban area or cluster, we assign them to the nearest urban area/cluster.

Cost-of-living information was collected roughly at the city level. CCER denotes these cities by assigning them the name of the Census-defined "urban area" that is located most closely to the city. In addition, CCER identifies the "Core-based Statistical Area" (CBSA) within which the city is located. First, we only include the quarterly cost-of-living composition measures for 1990-1992, 2002-2004. We then take four steps to match the cost of living to institutions. First, we match urban area name from the CCER data to the urban area in which the college is located, which we identified based on the latitude and longitude of the zipcode centroid in which the institution is located. Those institutions matching are then assigned the mean composite cost-of-living over the relevant time period. Second, for those not matching via urban area name, we then aggregate the cost of living over the CBSA and match via the CBSA code. There are, however, a number of CBSA that have no measures for cost of living, as well as institutions which are not located in a CBSA. Third, for those not matching via CBSA code, we match to the mean composite cost of living over the state. Finally, there are no cost-of-living measures for Maine (1992 and 2004) or Rhode Island (1992). Institutions in these states are assigned the mean composite over all New England states.

Appendix B: Replication and Extension of Long (2004)

To provide a direct comparison with previous work, we first extend the analysis of Long (2004) by including measures of college consumption amenities into her conditional logit specifications. In Table B1, the first two columns for each cohort year show her results (BTL) and our results (JMS) for a comparable specification side by side, indicating that we are able to successfully replicate her findings. It should be noted that our results should not be exactly comparable to hers since her estimation includes two-year colleges and students (which we exclude) and not all variables are interacted with sector in her model. The third column for each cohort adds three measures of consumption amenities to this basic model. We find that spending on student services and auxiliary enterprises have a large and statistically significant relationship with the likelihood of choosing a particular college, as does the presence of a division 1 basketball or football team and the fraction of students who join fraternities or sororities. The inclusion of these measures diminished the estimated importance of instructional expenditure.

Appendix C: Estimating the Probability of College Admissions

As noted in the Data section, both the NELS and ELS ask students to list colleges to which they applied and whether they were admitted to each college. We restrict our attention to student applications to the set of "regular" four-year colleges or universities in our main analysis sample. The resulting data set contains 22,934 (12,155) student-college observations from 2004 (1992). To determine the probability that individual i would be admitted to school j, we estimate probit models where the dependent variable is a binary indicator for admitted and the independent variables include student and school characteristics (and student x school interactions), including student race, gender, SES, high school GPA and standardized achievement scores along with measures of the school's selectivity such as the average SAT score of students in the school. Admissions models are estimated separately by the triple interaction of race, sector, and in-state status. Using the coefficient estimates from these models, we predict the likelihood that student i would be admitted to each of the college in our sample (regardless of whether or not the student actually applied to the college).

In order to separate admissions from enrollment decisions, we must first estimate the probability that student i would have been admitted to college j (conditional on applying).

NELS and ELS both ask students to report which colleges they applied to and, among these, to which colleges they were admitted. In the NELS, students were asked in 1992 (when they were high school seniors) to list up to 2 schools to which they had applied and to indicate whether or not they had been accepted to each school. In the 1994 follow-up survey, students were asked to list up to 5 schools they had attended since the 1992 survey. In order to capture a more complete set of schools to which the student may have applied, we combine information from both of these surveys. Specifically, we include all schools the student listed in the 1992 survey as well as the first two schools we observe the students attending based on the information reported in the 1994 survey (this survey provides enrollment dates which allow us to identify the first two schools). In this way, we observe a maximum of four application schools for each student. Also note that, by construction, a student will have been accepted to any school we observe him or her attending by 1994. Table C1 (C2) shows the distribution of applications and acceptances for the NELS (ELS) sample. Note that for this analysis we are incorporating information on all schools to which a student applied, including many two-year colleges that are not included in our analysis sample of colleges.

In the ELS, students were asked in 2006 (two years after expected high school graduation) to list up to 20 schools to which they applied, and whether they were accepted and/or attended. It also allows them to list the start and end dates of attendance.

Note that less than 0.1% of students listed the maximum possible number of schools in ELS, suggesting that we are capturing the full set of application schools for most students. In NELS, by contrast, over half of the students listed two different application schools in the 1992 survey, suggesting that even by including the extra information from the 1994 survey, we are likely missing at least some information on student application behavior.

We then estimate Probit models of the probability that student i was admitted to school j. In order to allow the admission function to vary across groups, we estimate separate models for each cohort year, and then within cohort year, we estimate separate models for six mutually exclusive set of student-school observations: 1) White or Asian students applying to in-state public colleges, 2) White or Asian students applying to out-state public colleges, 3) White or Asian students applying to private colleges, 4) other students applying to in-state public colleges, 5) other students applying to out-state public colleges, 6) other students applying to private colleges. We estimate separate models for racial minorities to allow for affirmative action policies. We estimate separate models for different school types to allow for admission preferences for in-state students in public universities and different admissions procedures in private colleges.

As predictors, we include several different measures of student academic ability, including high school GPA, 12th grade math score and the interaction between GPA and math score, a measure of student socioeconomic status, several measures of college selectivity, including the average SAT/ACT score of students in the college, the fraction of students admitted to the college and the log(enrollment). We also include a series of interactions between student ability and college selectivity. Finally, in order to allow for college preferences with regard to the geographic diversity of their students, we include a series of fixed effects for the region of the country in which the student went to high school (i.e., Northeast, South, Midwest, West) x the region in which the college is located.

Tables C3 and C4 show the results of these regressions. Because of the large number of higher-order terms and interactions, it is not productive to examine coefficients on specific predictors to assess the fit of the model. Instead, Tables C5 and C6 present summary statistics on the resulting predicted probabilities, broken out by various subgroups. The results all go in the expected direction and suggest that our predicted probabilities will provide good estimates.

Appendix D: Simulation-Based vs. Admissions Probability Weighted Estimates

Table D1 compares the predicted choice probability for each student-school observation implied by the weighting-based procedure and the predicted choice probability implied by the simulation-based procedure with 100 replications. Choice probabilities are calculated assuming the parameter estimates of specification (2) in Table 5. These correlations are also reported separately for student-school observations in each group defined by the predicted likelihood that the student would be accepted to that particular school.

Table A1: Number of institutions (starting with constructed sample)

	NELS (1992)	ELS (2004)
Total schools in sample	1,409	1457
No fallout for missing zip code, student or instructional spending	1,409	1457
After dropping schools with missing tuition or room and board costs	1,401	1452
No fallout for missing enrollment or mean SAT information	1,401	1452
After dropping schools that no student in sample chose	977	1108

Table A2: Summary Statistics on Sample Construction

Number of students	NELS (1992)	ELS (2004)
Total students in survey	28,622	16,197
After dropping students not enrolled in 12th grade at time of the 1992 or 2004 survey	17,959	13,370
After dropping students who did not respond to the follow-up survey	16,409	11,984
After dropping students who did not attend any postsecondary school within two years of expected high school graduation	8,571	9,466
After dropping students who did not attend a sample school	5,104	5,757
After dropping students with missing information on key covariates	4,101	5,757
After dropping students whose choice college was missing information	**4,083**	**5,741**

Table B1: College choice conditional on attendance

Dependent variable: attended college j within 2 years of high school graduation (odds ratios)

	1972			1980			1992		
	BTL	JMS	JMS	BTL	JMS	JMS	BTL	JMS	JMS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Tuition & Fees ($1000)	0.4686**	0.543***	0.523***	0.5809**	0.573***	0.555***	.6548**	0.755***	0.738***
	[32.32]	(0.00979)	(0.010)	[26.68]	(0.0114)	(0.0117)	[39.21]	(0.00639)	(0.00656)
Tuition & Fees ($1000) sq	1.0485**	1.030***	1.030***	1.0328**	1.030***	1.030***	1.0147**	1.008***	1.008***
	[24.87]	(0.00116)	(0.001)	[21.98]	(0.00128)	(0.00138)	[31.91]	(0.000292)	(0.000308)
Distance (100mi)	.1665**	0.213***	0.208***	0.1954**	0.204***	0.196***	.2668**	0.235***	0.235***
	[65.29]	(0.00522)	(0.005)	[60.91]	(0.00583)	(0.00571)	[64.66]	(0.00525)	(0.00526)
Instruct expend. ($1000)	1.038	1.053***	0.992	1.0303	1.071***	1.029	1.1035**	1.040***	1.023**
	[1.46]	(0.0200)	(0.019)	[1.27]	(0.0191)	(0.0184)	[6.08]	(0.00929)	(0.00992)
% Faculty with PhD	1.0050**	1.233***	1.106	1.0048**	0.950	0.903	1.0060**	1.266***	1.222***
	[7.18]	(0.0752)	(0.068)	[5.46]	(0.0656)	(0.0629)	[6.20]	(0.0918)	(0.0902)
Enrollment (100)	not	1.052***	1.048***	not	1.045***	1.040***	not	1.070***	1.062***
	reported	(0.00333)	(0.004)	reported	(0.00349)	(0.00417)	reported	(0.00401)	(0.00506)
Enrollment (100) sq	not	1.000***	1.000***	not	1.000***	1.000***	not	0.999***	1.000***
	reported	(4.71e-05)	(0.000)	reported	(4.33e-05)	(4.95e-05)	reported	(6.71e-05)	(7.56e-05)
Student - School test score ptile (pos)	0.6525**	0.805***	0.815***	0.8662**	0.858***	0.883***	.7129**	0.850***	0.875***
	[10.26]	(0.0285)	(0.029)	[4.64]	(0.0357)	(0.0370)	[11.26]	(0.0385)	(0.0398)
Student - School test score ptile (neg)	0.995	0.899***	0.898***	0.8324**	0.886***	0.885***	1.1809**	0.808***	0.784***
	[0.16]	(0.0338)	(0.034)	[5.75]	(0.0352)	(0.0352)	[4.78]	(0.0324)	(0.0317)
Student services + auxilary expend. ($1000)			1.457***			1.371***			1.209***
			(0.045)			(0.0496)			(0.0281)
Has Div1 Basketball/Football			1.202***			1.200***			1.212***
			(0.054)			(0.0629)			(0.0593)
% of Students who join Frat/Sor			2.421***			2.095***			2.052***
			(0.322)			(0.358)			(0.293)
Individuals	5,666			4881			5,693		
Observations	12,118,588	4,108,256	4,108,256	9,651,768	2,566,527	2,566,527	15,011,370	4,006,240	4,006,240

Notes: [z-statistics] or (standard errors) reported below odds ratio. *** p<0.01, ** p<0.05, * p<0.1. All specifications also include a square and cubic in distance, square in cost, expenditure squared, and student-school match variables squared. BTL does not interact % faculty with PhD or student-school match variables with sector (2-year or 4-year), so our estimates for 4-year college students only are not directly comparable.

Table C1: Application and Acceptance Rates for NELS Cohort (High School Class of 1992)

Panel A: Applications to 4-year institutions

Number of different 4-year institutions to which the student applied	Proportion of sample	Number of different 4-year institutions to which the student was accepted			
		1	2	3	4
1	35.7%	100.0%	0.0%	0.0%	0.0%
2	49.1%	40.4%	59.6%	0.0%	0.0%
3	13.9%	35.5%	34.6%	29.9%	0.0%
4	1.2%	0.0%	55.1%	32.7%	12.2%

Panel B: Applications to all institutions

Number of different institutions (any type) to which the student applied	Proportion of sample	Number of different institutions (any type) to which the student was accepted			
		1	2	3	4
1	27.9%	100.0%	0.0%	0.0%	0.0%
2	50.7%	37.0%	63.0%	0.0%	0.0%
3	19.0%	26.9%	38.5%	34.6%	0.0%
4	2.4%	0.0%	46.4%	29.9%	23.7%

Table C2: Application and Acceptance Rates for ELS (High School Class of 2004)

Restricted to eligible institutions		Number of schools accepted to									
Number of different schools applied to	Proportion of sample	1	2	3	4	5	6	7	8	9	10+
1	22.2%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	23.4%	25.0%	75.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	18.8%	11.7%	29.5%	58.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4	12.9%	6.6%	20.8%	27.1%	45.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5	8.9%	3.3%	11.4%	21.4%	30.8%	33.0%	0.0%	0.0%	0.0%	0.0%	0.0%
6	5.5%	1.9%	5.1%	17.1%	25.9%	25.9%	24.1%	0.0%	0.0%	0.0%	0.0%
7	3.5%	1.5%	11.3%	8.9%	17.7%	19.7%	20.7%	20.2%	0.0%	0.0%	0.0%
8	1.7%	2.0%	3.0%	5.1%	24.2%	17.2%	19.2%	20.2%	9.1%	0.0%	0.0%
9	1.3%	2.7%	1.3%	2.7%	8.0%	13.3%	20.0%	16.0%	24.0%	12.0%	0.0%
10+	1.8%	1.0%	3.0%	3.0%	7.9%	11.9%	19.8%	10.9%	14.9%	10.9%	16.8%

Unrestricted by institution		Number of schools accepted to									
Number of different schools applied to	Proportion of sample	1	2	3	4	5	6	7	8	9	10+
1	16.3%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	22.1%	18.6%	81.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	20.6%	9.2%	29.6%	61.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4	15.2%	4.5%	17.6%	28.3%	49.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5	10.1%	3.1%	10.7%	21.8%	30.1%	34.3%	0.0%	0.0%	0.0%	0.0%	0.0%
6	6.3%	1.9%	4.7%	14.2%	25.3%	26.4%	27.5%	0.0%	0.0%	0.0%	0.0%
7	3.9%	1.4%	9.0%	7.7%	21.2%	22.1%	20.7%	18.0%	0.0%	0.0%	0.0%
8	2.1%	1.7%	8.3%	18.2%	19.8%	18.2%	20.7%	13.2%	0.0%	0.0%	0.0%
9	1.5%	1.1%	2.3%	2.3%	6.8%	17.0%	18.2%	20.5%	20.5%	11.4%	0.0%
10+	1.9%	3.6%	5.5%	9.1%	13.6%	21.8%	15.5%	13.6%	10.9%	4.5%	1.8%

Table C3: Probit Estimation for Predicted Probabilities of Admission - NELS 1992

	White or Asian			All Others		
	Public		Private	Public		Private
	In State	Out of State	Both	In State	Out of State	Both
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	accepted	accepted	accepted	accepted	accepted	accepted
Student Grade Point Average	-0.391	4.473	-4.173	8.982**	5.948	-5.321
	(2.294)	(5.531)	(2.840)	(3.802)	(11.612)	(3.728)
Missing GPA	0.241**	0.579***	0.291***	0.012	-0.554	0.209
	(0.106)	(0.211)	(0.102)	(0.188)	(0.849)	(0.233)
Student Standardized Math Score	1.841**	0.964	0.613	1.061	-3.143	1.249
	(0.735)	(1.723)	(0.873)	(1.191)	(3.498)	(1.046)
Student Math Score * Student GPA	0.132	1.527**	-0.413	0.369	-0.473	0.428
	(0.248)	(0.600)	(0.277)	(0.615)	(2.229)	(0.522)
Student Socioeconomic Status	0.229	-0.133	0.026	-0.145	-0.643	0.173
	(0.211)	(0.452)	(0.227)	(0.333)	(0.993)	(0.398)
Log Enrollment	-0.060*	-0.051	0.018	-0.153**	0.036	-0.085
	(0.036)	(0.068)	(0.030)	(0.063)	(0.174)	(0.075)
Mean College SAT Score	-0.001	-0.054*	-0.024*	-0.022	-0.064	-0.012
	(0.014)	(0.031)	(0.014)	(0.024)	(0.092)	(0.022)
College Admission Rate	1.515	-2.562	-0.147	-1.035	1.174	1.758
	(1.325)	(3.195)	(1.559)	(2.159)	(6.538)	(2.256)
School Mean SAT * Student GPA	-0.011	-0.061	0.048	-0.124**	-0.257	0.044
	(0.030)	(0.063)	(0.031)	(0.052)	(0.213)	(0.047)
School Mean SAT * Student Math Score	-0.015	-0.005	-0.012	-0.009	0.059	-0.014
	(0.011)	(0.024)	(0.011)	(0.016)	(0.046)	(0.013)
School Mean SAT * Student Math * StudentGPA	-0.002	-0.009**	0.001	0.001	-0.018	-0.009*
	(0.002)	(0.005)	(0.002)	(0.004)	(0.021)	(0.004)
School Mean SAT * Student SES	-0.003	0.001	-0.000	-0.001	0.000	-0.001
	(0.002)	(0.004)	(0.002)	(0.003)	(0.007)	(0.003)
College Admission Rate * Student GPA	-0.861	-6.093	3.781	-11.391**	-15.309	6.204
	(2.885)	(6.749)	(3.589)	(4.628)	(14.996)	(4.823)
College Admission Rate * Student Math Score	-1.977**	-0.682	-0.113	-1.502	6.633	-1.472
	(0.870)	(2.093)	(1.101)	(1.485)	(4.779)	(1.372)
College Admission Rate * Student Math Score * Student GPA	-0.091	-1.146**	0.271	-0.486	2.232	0.166
	(0.205)	(0.505)	(0.236)	(0.588)	(1.599)	(0.550)
College Admission Rate * Student SES	-0.013	0.074	0.046	0.394	1.216	-0.021
	(0.164)	(0.330)	(0.154)	(0.307)	(0.946)	(0.348)
Squared Student GPA	0.663	-1.798	1.508	-3.911**	-2.412	2.326
	(1.006)	(2.641)	(1.294)	(1.722)	(5.596)	(1.677)
Squared Student Math Score	0.064	0.684	-0.363	-1.028	-0.444	-1.469
	(0.606)	(1.275)	(0.763)	(1.024)	(2.747)	(0.994)
School Mean SAT * College Admission Rate	-0.001	0.045	0.017	0.025	0.138	-0.005
	(0.018)	(0.039)	(0.019)	(0.030)	(0.127)	(0.030)
School Mean SAT * College Admission Rate * Student Math Score	0.021*	0.006	0.013	0.023	-0.111*	0.021
	(0.013)	(0.030)	(0.015)	(0.020)	(0.065)	(0.018)
School Mean SAT * College Admission Rate * Student GPA	0.028	0.062	-0.046	0.153**	0.486*	-0.047
	(0.039)	(0.078)	(0.040)	(0.067)	(0.291)	(0.064)
School Mean SAT * College Admission Rate * Squared Student GPA	-0.003	-0.017	0.011	-0.073**	-0.240*	0.023
	(0.017)	(0.037)	(0.018)	(0.030)	(0.143)	(0.028)
School Mean SAT * College Admission Rate * Squared Student Math Score	0.004	0.010	-0.010	-0.009	0.015	-0.017
	(0.010)	(0.019)	(0.011)	(0.017)	(0.056)	(0.016)
School Mean SAT * Squared Student GPA	-0.002	0.021	-0.015	0.057**	0.124	-0.019
	(0.013)	(0.029)	(0.014)	(0.024)	(0.104)	(0.021)
College Admission Rate * Squared Student GPA	-0.253	2.330	-1.092	5.125**	6.982	-2.749
	(1.252)	(3.226)	(1.629)	(2.088)	(7.283)	(2.173)
School Mean SAT * Squared Student Math Score	-0.001	-0.007	0.008	0.011	-0.006	0.019
	(0.008)	(0.016)	(0.009)	(0.013)	(0.040)	(0.012)
College Admission Rate * Squared Student Math Score	-0.271	-0.927	0.358	0.965	0.568	1.389
	(0.721)	(1.524)	(0.968)	(1.260)	(3.739)	(1.277)
Constant	-0.394	3.544	0.952	2.657	-2.507	0.412
	(1.077)	(2.657)	(1.243)	(1.798)	(5.187)	(1.815)
Observations	4990	1335	3873	1124	212	621

Coefficient Estimates, Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Variables included but not presented: Interactions between Student and Institution Census (4) Region.

Out-of-State specification (2,3,5,6), combine non-Northeast Institution region to account for relatively few students traveling to out-of-state public schools

Table C4: Probit Estimation for Predicted Probabilities of Admission - ELS 2004

	White or Asian			All Others		
	Public		Private	Public		Private
	In State	Out of State	Both	In State	Out of State	Both
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	accepted	accepted	accepted	accepted	accepted	accepted
Student Grade Point Average	1.115**	4.407***	1.717**	-0.054	0.634	0.923
	(0.560)	(1.668)	(0.753)	(0.464)	(0.933)	(0.783)
Missing GPA	0.393***	0.373***	0.229***	0.060	0.071	0.140
	(0.079)	(0.135)	(0.080)	(0.094)	(0.199)	(0.128)
Student Standardized Math Score	0.596	0.519	0.595	1.270***	-0.682	-0.163
	(0.481)	(1.516)	(0.679)	(0.459)	(0.890)	(0.764)
Student Math Score * Student GPA	-0.514	-0.052	-0.170	-0.118	-1.194**	0.823*
	(0.317)	(0.863)	(0.439)	(0.329)	(0.564)	(0.435)
Student Socioeconomic Status	0.044	-0.711*	0.084	0.162	0.092	0.131
	(0.136)	(0.395)	(0.229)	(0.154)	(0.291)	(0.238)
Log Enrollment	-0.129***	-0.125*	-0.081***	-0.093**	-0.103	-0.115**
	(0.037)	(0.064)	(0.028)	(0.046)	(0.083)	(0.047)
Mean College SAT Score	-0.039***	-0.028	-0.043***	-0.046***	-0.026**	-0.041***
	(0.008)	(0.019)	(0.009)	(0.007)	(0.012)	(0.008)
College Admission Rate	1.241*	0.749	0.558	-0.789	-0.184	-1.434
	(0.723)	(1.828)	(0.951)	(0.649)	(1.193)	(0.909)
School Mean SAT * Student GPA	-0.012	-0.053**	-0.026**	0.005	-0.015	-0.005
	(0.009)	(0.024)	(0.010)	(0.007)	(0.014)	(0.010)
School Mean SAT * Student Math Score	-0.007	-0.010	-0.007	-0.013*	0.015	0.006
	(0.008)	(0.022)	(0.009)	(0.007)	(0.013)	(0.009)
School Mean SAT * Student Math * StudentGPA	0.005	0.001	0.002	-0.003	0.014***	-0.003
	(0.003)	(0.008)	(0.004)	(0.003)	(0.005)	(0.004)
School Mean SAT * Student SES	-0.002	0.004	0.000	-0.001	0.003	-0.001
	(0.001)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)
College Admission Rate * Student GPA	-1.021	-4.985**	-2.131**	0.278	-0.666	-0.465
	(0.803)	(2.182)	(0.998)	(0.696)	(1.420)	(1.121)
College Admission Rate * Student Math Score	-0.202	0.071	-0.406	-1.216*	1.432	1.045
	(0.700)	(1.956)	(0.904)	(0.695)	(1.349)	(1.096)
College Admission Rate * Student Math Score * Student GPA	0.288	-0.158	0.031	0.417	0.512	-0.709*
	(0.272)	(0.627)	(0.300)	(0.343)	(0.685)	(0.419)
College Admission Rate * Student SES	0.268**	0.739***	0.092	-0.012	-0.162	0.177
	(0.118)	(0.281)	(0.148)	(0.152)	(0.339)	(0.202)
Squared Student GPA	-0.725*	-0.111	0.216	0.013	1.119*	0.363
	(0.415)	(1.449)	(0.595)	(0.317)	(0.584)	(0.508)
Squared Student Math Score	-0.168	-1.544	0.699	-0.309	-0.023	-0.794
	(0.352)	(1.007)	(0.507)	(0.388)	(0.725)	(0.537)
School Mean SAT * College Admission Rate	0.013	0.011	0.020	0.040***	0.018	0.031**
	(0.012)	(0.025)	(0.012)	(0.012)	(0.021)	(0.012)
School Mean SAT * College Admission Rate * Student Math Score	0.008	0.010	0.012	0.018	-0.020	-0.013
	(0.012)	(0.029)	(0.012)	(0.011)	(0.020)	(0.014)
School Mean SAT * College Admission Rate * Student GPA	0.021	0.074**	0.041***	0.002	0.024	0.004
	(0.014)	(0.033)	(0.014)	(0.012)	(0.023)	(0.015)
School Mean SAT * College Admission Rate * Squared Student GPA	-0.020**	-0.013	0.003	-0.002	0.020	0.012
	(0.010)	(0.023)	(0.009)	(0.008)	(0.014)	(0.009)
School Mean SAT * College Admission Rate * Squared Student Math Score	0.002	-0.022	0.009	-0.005	-0.008	-0.003
	(0.008)	(0.016)	(0.007)	(0.009)	(0.015)	(0.009)
School Mean SAT * Squared Student GPA	0.019***	0.008	0.003	0.005	-0.015*	-0.009
	(0.006)	(0.018)	(0.007)	(0.005)	(0.008)	(0.006)
College Admission Rate * Squared Student GPA	0.654	0.292	-0.711	-0.133	-1.510*	-0.442
	(0.595)	(1.780)	(0.738)	(0.465)	(0.908)	(0.702)
School Mean SAT * Squared Student Math Score	0.004	0.020	-0.005	0.006	0.001	0.004
	(0.005)	(0.012)	(0.006)	(0.006)	(0.010)	(0.006)
College Admission Rate * Squared Student Math Score	-0.250	1.695	-1.060*	0.181	0.582	0.921
	(0.508)	(1.253)	(0.623)	(0.572)	(1.093)	(0.777)
Constant	2.663***	2.291	2.603***	2.907***	1.843**	3.608***
	(0.535)	(1.465)	(0.722)	(0.509)	(0.933)	(0.697)
Observations	7893	2175	6733	3266	826	2041

Coefficient Estimates, Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Variables included but not presented: Interactions between Student and Institution Census (4) Region.
Specification (5), however, combines non-Northeast Institution regions to account for relatively few minority students traveling to out-of-state public schools

Table C5: Summary Statistics for Predicted Probability of Admission - NELS 1992 only

			Percentile					
	min	10th	50th	90th	max	mean	Unique Students	Stu x Sch
White/Asian, Public, In-state School	7.5%	59.5%	80.0%	89.4%	99.3%	76.9%	7,478	6,110
White/Asian, Public, Out-of-state School	0.0%	50.5%	77.9%	89.6%	100.0%	73.8%		1,682
White/Asian, Private School	8.2%	46.1%	74.0%	88.9%	99.9%	70.5%		4,848
Minority, Public, In-state School	10.0%	45.5%	70.7%	89.5%	99.6%	68.8%	1,469	1,370
Minority, Public, Out-of-state School	0.0%	24.2%	74.5%	99.5%	100.0%	68.4%		284
Minority, Private School	0.8%	34.5%	60.6%	79.0%	99.7%	58.6%		759
All Races and Schools: Rank by Quartile								
Top Students by Standardized Math Score	0.0%	63.1%	83.3%	92.6%	100.0%	79.8%	1,022	102,200
Bottom Students by Standardized Math Score	0.0%	32.5%	58.9%	79.8%	100.0%	57.3%	1,022	102,200
Top Students by Standardized Grade 12 GPA	0.0%	40.4%	71.5%	92.4%	100.0%	68.4%	1,265	126,500
Bottom Students by Standardized Grade 12 GPA	0.0%	39.8%	69.8%	87.3%	100.0%	66.5%	1,043	104,300
Top Schools by Mean SAT Score	0.0%	28.8%	57.8%	76.8%	100.0%	55.1%	1,038	102,572
Bottom Schools by Mean SAT Score	0.0%	65.1%	84.8%	93.5%	100.0%	81.2%	1,172	117,444
Top Schools by Lowest Admit Rate	0.0%	27.7%	59.1%	83.6%	100.0%	57.2%	1,038	101,734
Bottom Schools by Lowest Admit Rate	0.0%	60.5%	82.1%	92.1%	100.0%	78.6%	1,028	102,781
Top Students (Math) and Top Schools (SAT)	0.0%	41.5%	68.8%	81.3%	100.0%	65.1%	269	25,823
Top Students (Math) and Bottom Schools (SAT)	10.5%	78.2%	89.7%	96.2%	100.0%	88.3%	286	29,289
Bottom Students (Math) and Top Schools (SAT)	0.0%	23.9%	42.2%	64.5%	100.0%	43.7%	257	25,564
Bottom Students (Math) and Bottom Schools (SAT)	0.0%	48.3%	71.6%	84.5%	100.0%	68.3%	280	29,515

Table C6: Summary Statistics for Predicted Probability of Admission - ELS 2004 only

	Percentile							
	min	10th	50th	90th	max	mean	Unique Students	Stu x Sch
White/Asian, Public, In-state School	3.8%	59.8%	85.1%	93.7%	99.5%	80.2%	5,737	7,893
White/Asian, Public, Out-of-state School	0.0%	55.4%	86.4%	96.1%	100.0%	80.5%		2,175
White/Asian, Private School	0.3%	41.2%	88.0%	95.9%	100.0%	78.3%		6,733
Minority, Public, In-state School	3.8%	38.5%	74.7%	89.5%	99.7%	69.1%	2,127	3,266
Minority, Public, Out-of-state School	1.8%	39.2%	69.8%	90.3%	100.0%	67.0%		826
Minority, Private School	0.0%	36.1%	78.1%	97.5%	100.0%	71.8%		2,041
All Races and Schools: Rank by Quartile								
Top Students by Standardized Math Score	0.0%	71.3%	91.7%	97.9%	100.0%	86.9%	1,437	143,700
Bottom Students by Standardized Math Score	0.0%	27.5%	66.3%	89.3%	100.0%	62.2%	1,436	143,600
Top Students by Standardized Grade 12 GPA	0.7%	76.2%	92.7%	98.3%	100.0%	88.5%	1,469	146,900
Bottom Students by Standardized Grade 12 GPA	0.0%	27.5%	66.3%	89.2%	100.0%	62.1%	1,477	147,700
Top Schools by Mean SAT Score	0.0%	17.8%	67.8%	95.1%	100.0%	61.9%	1,544	148,711
Bottom Schools by Mean SAT Score	0.0%	67.4%	89.2%	97.6%	100.0%	85.1%	1,411	146,656
Top Schools by Lowest Admit Rate	0.0%	17.0%	68.4%	96.8%	100.0%	62.4%	1,413	143,674
Bottom Schools by Lowest Admit Rate	0.0%	65.4%	88.2%	96.0%	100.0%	84.0%	1,426	142,875
Top Students (Math) and Top Schools (SAT)	0.0%	39.3%	88.2%	98.1%	100.0%	78.3%	368	37,545
Top Students (Math) and Bottom Schools (SAT)	0.3%	76.2%	90.8%	98.7%	100.0%	88.4%	378	36,708
Bottom Students (Math) and Top Schools (SAT)	0.0%	10.9%	42.4%	76.8%	98.0%	43.3%	398	37,025
Bottom Students (Math) and Bottom Schools (SAT)	0.0%	50.5%	80.0%	93.4%	100.0%	75.4%	333	37,049

Table D1. Comparison between Weighted and Simulation-Based Likelihood Value

Estimated prob(admit) (rounded)	Sample in Group: N	Fraction of sample	Cumulative Fraction	Weighted-Simulation Correlation
0.00	22,419	0.22	0.22	0.7306
0.05	48,773	0.47	0.69	0.8887
0.10	57,569	0.56	1.24	0.7410
0.15	74,253	0.72	1.96	0.8624
0.20	91,996	0.89	2.85	0.9154
0.25	109,743	1.06	3.91	0.8857
0.30	133,276	1.29	5.20	0.9396
0.35	165,199	1.60	6.79	0.9454
0.40	194,212	1.88	8.67	0.9526
0.45	234,775	2.27	10.94	0.9677
0.50	284,237	2.75	13.69	0.9756
0.55	349,685	3.38	17.06	0.9831
0.60	431,761	4.17	21.24	0.9843
0.65	530,792	5.13	26.36	0.9900
0.70	656,128	6.34	32.70	0.9907
0.75	758,903	7.33	40.04	0.9941
0.80	889,011	8.59	48.62	0.9961
0.85	1,054,220	10.19	58.81	0.9975
0.90	1,214,993	11.74	70.55	0.9984
0.95	1,707,180	16.49	87.04	0.9994
1.00	1,340,990	12.96	100.00	0.9997
All	10,350,115	100		0.9879

Notes: Probability of admission is estimated flexibly as described in Appendix C. Final column depicts the simple correlation between the predicted choice probability implied by the weighting-based procedure and the predicted choice probability implied by the simulation-based procedure with 100 replications. Choice probabilities are calculated assuming the parameter estimates of specification (2) in Table 5. These correlations are reported separately for individual-school observations in each admissions probability group.

Table E1: OLS Estimation for Predicted Price Ratios - NPSAS

	1996						2004					
	White or Asian			All Others			White or Asian			All Others		
	Public		Private	Public		Private	Public		Private	Public		Private
	In State	Out of State	Both	In State	Out of State	Both	In State	Out of State	Both	In State	Out of State	Both
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio	price_ratio
Student Standardized Math Score	-0.166***	-0.287***	-0.184***	-0.121	-0.0718	-0.199***	-0.139***	-0.342***	-0.159***	-0.108	0.0381	0.0510
	(0.0327)	(0.0669)	(0.0378)	(0.0744)	(0.208)	(0.0632)	(0.0455)	(0.124)	(0.0390)	(0.0753)	(0.239)	(0.0676)
Student Standardized Income	0.194***	0.118**	0.278***	0.234***	0.0650	0.259***	0.194***	-0.0717	0.140***	0.412***	0.0669	0.104*
	(0.0321)	(0.0516)	(0.0327)	(0.0887)	(0.194)	(0.0718)	(0.0387)	(0.0862)	(0.0260)	(0.0778)	(0.202)	(0.0556)
Mean College SAT Score	0.0103***	0.0148**	-0.00677*	0.00480	0.0324*	0.0134***	0.0184***	0.0162	-0.0145***	0.00941*	0.0156	-0.00722*
	(0.00280)	(0.00646)	(0.00354)	(0.00475)	(0.0171)	(0.00388)	(0.00461)	(0.0104)	(0.00308)	(0.00495)	(0.0170)	(0.00408)
Student Math * School Mean SAT	0.00160***	0.00364***	0.00157***	0.000982	0.00204	0.00268***	0.00102	0.00434**	0.00120*	0.000731	0.000959	-0.00104
	(0.000510)	(0.00101)	(0.000526)	(0.00102)	(0.00291)	(0.000841)	(0.000760)	(0.00208)	(0.000614)	(0.00116)	(0.00361)	(0.000974)
Student Math * Student Income	-0.0364***	0.00668	0.000529	-0.0884***	-0.113*	-0.00275	-0.0337***	-0.0207	-0.00608	0.0137	-0.0434	-0.0228
	(0.00806)	(0.0151)	(0.00610)	(0.0295)	(0.0621)	(0.0221)	(0.00955)	(0.0200)	(0.00636)	(0.0257)	(0.0790)	(0.0167)
School Mean SAT * Student Income	0.0000422	0.000224	-0.00161***	0.00134	0.00258	-0.000210	0.00119*	0.00238*	0.000158	-0.00142	0.00253	0.00158*
	(0.000493)	(0.000758)	(0.000415)	(0.00120)	(0.00284)	(0.000976)	(0.000629)	(0.00138)	(0.000376)	(0.00122)	(0.00326)	(0.000839)
Squared Student Math Score	-0.0309***	-0.0594***	-0.0150**	-0.00681	0.0164	-0.00533	-0.0217***	-0.0353*	-0.0172**	-0.0168	0.0219	0.0245*
	(0.00573)	(0.0128)	(0.00593)	(0.0133)	(0.0396)	(0.0140)	(0.00741)	(0.0197)	(0.00727)	(0.0155)	(0.0710)	(0.0129)
Squared Student Income	-0.00597***	-0.00829***	-0.00406***	-0.0215***	-0.0374	-0.0202***	-0.0420***	-0.00830	-0.0160***	-0.0469***	-0.0638	-0.0277***
	(0.000442)	(0.00127)	(0.000314)	(0.00345)	(0.0301)	(0.00563)	(0.00341)	(0.00654)	(0.00168)	(0.00719)	(0.0474)	(0.00437)
Squared School Mean SAT	-8.42e-05***	-9.02e-05*	7.31e-05***	-4.62e-05	-0.000260**	-0.000117***	-0.000164***	-0.000124	0.000148***	-0.000115***	-0.000164	5.65e-05*
	(2.24e-05)	(5.03e-05)	(2.55e-05)	(3.68e-05)	(0.000130)	(3.05e-05)	(3.89e-05)	(8.96e-05)	(2.45e-05)	(4.18e-05)	(0.000160)	(3.34e-05)
Constant	0.336***	0.175	0.641***	0.346**	-0.250	0.132	0.0535	0.233	0.757***	0.282*	0.377	0.620***
	(0.0847)	(0.203)	(0.120)	(0.154)	(0.531)	(0.121)	(0.134)	(0.301)	(0.0948)	(0.149)	(0.438)	(0.123)
Observations	4679	781	3436	1029	121	646	2886	384	2507	788	87	569

Coefficient Estimates, Standard errors in parenthe

*** p<0.01, ** p<0.05, * p<0.1

TableE2: Summary Statistics for Predicted Price Ratio - NELS 1992 and ELS 2004

	1992						2004					
		Percentile						Percentile				
	min	10th	50th	90th	max	mean	min	10th	50th	90th	max	mean
White/Asian, Public, In-state School	0.0%	41.1%	61.9%	81.3%	100.0%	61.5%	0.0%	21.9%	53.5%	75.7%	100.0%	50.9%
White/Asian, Public, Out-of-state School	0.0%	42.5%	68.2%	86.9%	100.0%	66.0%	0.0%	32.0%	62.4%	80.0%	98.1%	58.8%
White/Asian, Private School	0.0%	34.2%	57.0%	74.7%	100.0%	55.5%	0.0%	25.1%	45.5%	67.5%	100.0%	45.8%
Minority, Public, In-state School	0.0%	0.0%	39.0%	71.9%	100.0%	37.6%	0.0%	0.0%	42.7%	78.0%	100.0%	40.5%
Minority, Public, Out-of-state School	0.0%	18.1%	68.4%	92.0%	100.0%	61.8%	0.0%	30.9%	74.3%	93.0%	100.0%	68.0%
Minority, Private School	0.0%	1.6%	42.9%	72.7%	100.0%	40.8%	0.0%	11.3%	45.0%	63.8%	96.6%	41.5%
All Races and Schools: Rank by Quartile												
Top Students by Standardized Math Score	0.0%	22.9%	48.3%	73.5%	100.0%	48.4%	0.0%	22.7%	48.2%	73.4%	100.0%	48.3%
Bottom Students by Standardized Math Score	0.0%	21.4%	59.8%	82.9%	100.0%	56.0%	0.0%	22.0%	59.8%	82.9%	100.0%	56.1%
Top Students by Standardized Grade 12 GPA	0.0%	25.5%	53.8%	79.1%	100.0%	52.8%	0.0%	23.4%	51.3%	77.0%	100.0%	50.6%
Bottom Students by Standardized Grade 12 GPA	0.0%	27.0%	59.0%	82.1%	100.0%	56.4%	0.0%	29.1%	58.7%	81.5%	100.0%	56.6%
Top Schools by Mean SAT Score	0.0%	36.5%	64.7%	83.3%	100.0%	61.7%	0.0%	33.5%	61.9%	82.3%	100.0%	59.3%
Bottom Schools by Mean SAT Score	0.0%	22.0%	49.3%	74.7%	100.0%	48.7%	0.0%	19.8%	47.3%	73.2%	100.0%	46.9%
Top Students (Math) and Top Schools (SAT)	0.0%	45.2%	65.2%	81.3%	100.0%	64.0%	0.0%	39.0%	61.4%	79.2%	100.0%	60.1%
Top Students (Math) and Bottom Schools (SAT)	0.0%	11.4%	35.7%	60.8%	100.0%	36.4%	0.0%	7.0%	32.6%	58.7%	100.0%	33.2%
Bottom Students (Math) and Top Schools (SAT)	0.0%	12.1%	60.7%	82.8%	100.0%	54.9%	0.0%	15.1%	59.3%	82.4%	100.0%	54.4%
Bottom Students (Math) and Bottom Schools (SAT)	0.0%	26.4%	61.0%	80.9%	100.0%	57.2%	0.0%	27.4%	60.3%	79.3%	100.0%	56.6%