# DO AND SHOULD FINANCIAL AID PACKAGES AFFECT STUDENTS' COLLEGE CHOICES?

Christopher Avery

and

Caroline M. Hoxby

Every year, thousands of high school seniors with high college aptitude face complicated "menus" of scholarship and aid packages designed to affect their college choices. Using an original survey designed for this paper, we investigate whether students respond to their "menus" like rational human capital investors. Whether they make the investments efficiently is important not only because they are the equivalent of the "Fortune 500" for human capital, but also because they are likely to be the most analytic and long-sighted student investors. We find that the typical high aptitude student chooses his college and responds to aid in a manner that is broadly consistent with rational investment. However, we also find some serious anomalies: excessive response to loans and work-study, strong response to superficial aspects of a grant (such as whether it has a name), and response to a grant's share of college costs rather than its amount. Approximately 30 percent of high aptitude students respond to aid in a way that apparently reduces their lifetime present value. While both a lack of sophistication/information and credit constraints can explain the behavior of this 30 percent of students, the weight of the evidence favors a lack of sophistication.

The authors are affiliated, respectively, with the John F.Kennedy School of Government and the Department of Economics, Harvard University. The authors gratefully thank Andrew Fairbanks and Jim Barker, who helped to design and implement the College Admissions Project survey. 510 high school counselors made the survey happen; the authors wish they could thank them individually. The authors owe thanks to the research assistants who contacted high school counselors, coded data, programmed, and otherwise regarded the data thoughtfully. Their perspicacity is what made the College Admissions Project data as interpretable and error-free as we are grateful to think they are: Joshua Barro, Rohit Chandwani, Michael Cuthbert, Suzanne Ko, Ilyana Kuziemko, Michael McNabb, Kathryn Markham, Emily Oster, Jenna Robins, Aaron Roth, Maria Shim, Catherine So, Rania Succar, Michael Thakur, and Kenneth Wang. Scott Resnick deserves special thanks.

## I. How Scholarships and Aid Affect the College Choices of High Aptitude Students

Every year, thousands of high school seniors who have high college aptitude are faced with complicated arrays of scholarships and aid packages that are intended to influence their college choices. Some of the scholarships and aid are meant purely to relieve liquidity constraints that might prevent needy students from attending the college they most prefer. Other scholarships and aid packages are designed to alter students' preference ranking of colleges–for instance, by attracting them to a college that might be unappealing in the absence of a scholarship. A student with high aptitude has complex financial supports for his college education: outside scholarships that are purely merit-based; outside scholarships that are merit- and need-based; state scholarships that are merit-based (and/or need-based) but usable only at in-state *public* colleges; state scholarships that are merit-based (and/or need-based) and usable at any in-state college; college scholarships that are purely merit-based; college scholarships that are merit- and need-based; college grants (as opposed to named scholarships) that are merit- and need-based; subsidized and unsubsidized loans from their college, outside charitable organizations, and the government; and work-study programs that subsidize on-campus work.

This fascinating array of scholarships, grants, loans, and work-study programs exists because many parties want to alter meritorious students' college choices. The parties' objectives are diverse–from a purely altruistic desire to relax constraints facing the needy to a college's self-interested desire to enroll high aptitude students who raise its profile or improve education for other students on campus.

The students who face these complex choices are an important, though small group. Many commentators would say that they are important because they will later account for a disproportionate share of the nation's leaders, scientists, and intellectuals. Their human capital and abilities are often thought to generate social spillovers. However, the behavior of high aptitude student is also important for purely scientific inquiry into human capital investment. They are capable of the largest human capital investments in the nation; by the time these students complete their education, some of them will be "walking capital stocks" of considerable size (income-generating power). In this era in which the *human* capital stock of developed economies like America's is thought to be crucial to growth, it is important to know whether the biggest investors in human capital make their investments efficiently. It is not only the size of their investments that makes them interesting: observing them allows us to witness the forces that affect human capital investments at their most highly charged because the stakes are high. Also, high aptitude students are likely to be the investors who most closely obey the model of rational human capital investor: they are capable of complex analysis, they are the least risky for creditors, and they tend to be patient people who take future benefits seriously.

Despite the interest inherent in the question of how meritorious students respond to scholarships and aid, very little evidence exists to answer the question. We believe that this is primarily because analysis is impossible

with traditional sources of student survey data, which do not contain sufficiently large numbers of this relatively rare type of student. That is, one cannot hope to use survey data to understand such students' behavior unless the survey greatly oversamples them. For this paper, we created a survey directed specifically to high aptitude students, with the result that we use the largest existing data set on this type of student.

Although almost no systematic evidence exists on how high aptitude students respond to scholarships and aid, many selective colleges do perform internal analyses using data on the students they themselves admit. While we believe that these analyses provide helpful evidence, they do have some flaws: they necessarily focus on a narrow set of students (the students they accept); they are often tacit about their empirical methods (some of which would not bear scrutiny); and the studies are hard to compare systemically because they are rarely published (their distribution is often purely internal). Also, while colleges have complete information about their own aid offers and matriculation, they often lack full information about where students who turn down their offers *do* matriculate and what aid they receive there. A notable exception, which is published, is Ehrenberg and Sherman's (1984) study of students who were accepted by Cornell University in the spring of 1981.

Though there is scarcity of systematic evidence on the college choice behavior of meritorious students who can attract complex offers of financial support, there is no similar scarcity about the effects of financial aid on the typical student or the poor student. We shall not attempt a survey here, but instead direct readers to the chapters by Avery and Kane (forthcoming), Dynarski (forthcoming), and Long (forthcoming) in this volume.

## II. How *Should* Students Respond to Scholarships and Aid?
## A Swift Review of the Standard Model of Human Capital Investment

Throughout this paper, we are working from a model of human capital investment. It sets the standard we use in our attempt to determine whether students react too much or too little to scholarships and aid. Because it underlies our question, a quick review seems in order.

It may be useful to state the implication of the model in intuitive terms. In return for getting more aid, a student must generally accept a reduction in the human capital investment made in him at college or a reduction in the consumption he enjoys at college. Put more bluntly, a student must generally enjoy a less resource-rich college environment or a less rich peer environment in return for larger grants and other subsidies.

A simple version of the human capital investment problem will show why students generally face these trade-offs. Consider the problem facing a student who has very high college aptitude. In the United States, it is reasonable to assume that he knows that he is going to attend *some* four-year college and must only decide which

college to attend among those that have admitted him.[1] If he acts as a rational investor, not bound by credit constraints (an issue we will consider below), then he need make only two calculations for each college in his choice set. Supposing that the student has figured out the cheapest way to attend each college, given the aid offered him, his first calculation is the present discounted cost of attending each college *j*:

$$\sum_{t=1}^{t=4} \frac{(TuitionFees_{jt} + RoomBoard_{jt})}{(1+\delta)^{t-1}} - \sum_{t=1}^{t=4} \frac{ApplicableGrants_{ijt} + ApplicableLoans^{*}_{ijt} + WorkStudySubsidies^{*}_{ijt})}{(1+\delta)^{t-1}} + \sum_{t=1}^{t=T} \frac{LoanRepayments^{*}_{ijt}}{(1+\delta)^{t-1}} . \quad (1)$$

His second calculation is the presented discounted value of the consumption he enjoys at college *j* plus the presented discounted value of the stream of income generated by the human capital invested in him at college *j*:

$$\sum_{t=1}^{t=4} \frac{FoodConsumption_{ijt} + HousingConsumption_{ijt} + OtherConsumption_{ijt}}{(1+\delta)^{t-1}} \sum_{t=5}^{t=T_i} \frac{r_{it}(Resources_{jt} + PeerSpillovers_{jt})}{(1+\delta)^{t-1}} . \quad (2)$$

In both equations (1) and (2), *i* indexes individual students, *j* indexes colleges, $\delta$ is the discount rate on future years, and *t* indexes years (*t=1* is the freshman year, *t=5* is the first post-baccalaureate year, and $T_i$ is the end of life).

In equation (1), the first term is the present discounted value of total potential costs of college: tuition, fees, room, and board. Notice that these costs apply only to the four years of college. The second term is the present discounted value of the potential costs that he does *not* (immediately) pay: the grants that apply to college *j* (college *j*'s institutional grants and outside scholarships usable at college *j*), the loans that apply to college *j* (college *j*'s institutional loans, subsidized loans from the federal government and outside charitable groups, and unsubsidized

[1] When he is applying to colleges, the student must form expectations of his probability of admission to each college and the scholarships or aid each college would likely offer him. That is, in order to avoid the inconvenience and cost of applying to all colleges, a student attempts to foresee the choices he will have and the actions he will take in the stage upon which we focus: the stage at which the student chooses among colleges that have accepted him. While we do not model the application stage because it is not necessary for our analysis, the extension of our model to the earlier stage requires only simple adaptations: application must have a cost (at least an effort or psychic cost, if not a financial one); students must use expected probabilities of admission; and students must use expected grants, loans, *et cetera*.

bank loans), and the subsidy value of the work-study program given the number of hours and job he works. The third term records the present discounted value of the payments the student makes (up to the end of his life if necessary!) in order to repay the college loans recorded in the second term. The variables that have asterisks require the student to choose them optimally.[2]

In equation (2), the first term is the present discounted value of the consumption that the students enjoys at the college: food, housing, and other consumption like recreational facilities, concerts, and so on. Of course, this consumption does not include consumption for which the student pays out of his own pocket, although it may include college-financed discounts at on-campus restaurants, concerts, *et cetera*. The second term is the present discounted value of the returns he enjoys on the human capital invested in him at college *j*. This human capital is assumed to come from two sources: resources invested in his learning (faculty time, college advisors' time, library resources, laboratories, *et cetera*) and knowledge spillovers from his peers. Peer spillovers are only a *possible* source of human capital; though they are generally believed to exist, their form and even their existence is somewhat doubtful. Good peers may merely facilitate a student's absorbing the non-peer resources invested in him; in this case the functional form, though not the spirit, of equation (2) is wrong. Notice that we allow the student's return on the stock of human capital he acquires in college to be specific to him (presumably because of ability) and specific to each year (presumably because human capital acquired at college interacts with human capital acquired through experience). The student's return on his stock of human capital need not be exclusively financial. Any return–psychic, social, *et cetera*–that the student values may be included in $r_{it}$. Of course, we will have difficulty quantifying non-financial returns.

The alert reader may notice that we have said nothing about the opportunity cost of college, which is the sum of the income and value of leisure the student sacrifices when he attends college. These opportunity costs are approximately the same for any college chosen by the meritorious student, so we do not need to consider them when we explain his college choice.

In order to the choose which college to attend, the student has only to subtract equation (1) from equation

---

[2] The optimal use of loans on offer should take into account each loan's interest rate, its repayment schedule, and its provisions (if any) in case of disability or other exogenous reason why the student might fail to have sufficient income to repay. In practice, the optimal use of loans tends to be simple: students exhaust more subsidized loans, then exhaust somewhat subsidized loans, leaving their marginal loan an unsubsidized one. The choice of optimal work-study hours is a good deal more complicated. A student should take into account the per-hour subsidy implicit in the work-study program and the loss of human capital caused by using hours for work that might be used for study. Note the subsidy he needs is the *true* value of the subsidy, compared to the market wage for an equivalent job. In order to consider an equivalent job, he will generally need to think about the equalizing wage differentials associated with sort of job provided under the work-study program–is it menial, educational, or in a particularly convenient location? We can observe very little of the information that we would need to assess the true subsidy value of work-study or to determine whether the student is choosing his work-study hours optimally. Therefore, we will have little to say about optimal work study after this.

(2) and consider the difference he obtains for each college. He should attend the college with the largest difference–that is, the college at which the present discounted benefits of college most exceed the present discounted costs.

It should now be clear that it could not be an equilibrium for students *not* to face trade-offs between aid and the resources available at a college (including peers). For instance, suppose a given student could be admitted to colleges A, B, and C, and that college C was preferable or at least as good as the others on the grounds of selectivity, the resources available for students, tuition (that is, lower tuition), campus life, location, and so on. Suppose also that college C systemically offered more aid (that is, systemically offered aid packages containing subsidies with a greater total value). Then, no trade-off would exist; the student would do better all around by matriculating at college C.

This no-trade off situation could not be an equilibrium. If all students saw the clear advantages of college C and received more aid, college C would be so over-subscribed that it would automatically become more selective, so that the typical student admitted to colleges A and B would no longer be admitted to college C. Of course, it is possible that an individual student will face no trade-off between two colleges in his choice set. Such no-trade-off situations must, however, be idiosyncratic to individual students and cannot hold generally.

We would say that a student is acting like a rational human capital investor if he always chooses the college that maximizes the difference between equation (2) and equation (1) for him. That is, he should never be tempted by more aid to attend a college that offers such reduced consumption and human capital investment that he is worse off, over his lifetime. Similarly, he should not refuse to attend colleges that offer aid packages that are so generous that they more than offset the reduction in consumption and human capital investment he experiences in college. Also, he should act in accordance with the presented discounted value of various forms of aid, recognizing that loans must be repaid and that only part of a work-study package is a subsidy. In this paper, when we test students against a standard of rational human capital investment, we are attempting to determine whether they act in accordance with the "should"s of the last few sentences.

There are three broad reasons why students might fail to respond to aid like the rational human capital investor. First, a student may be rational but credit constrained. In particular, his parents may be too well-off to attract need-based aid *and* unwilling to pay for the optimal college themselves *and* unwilling to co-sign loans so that he can pay for the optimal college himself. Second, a student may be rational but systemically misinformed–for instance, he may be naive about colleges' different levels of resources and therefore choose a college at which he will accumulate much less human capital than he thought he would. Third, a student may simply be irrational–that is, he may not maximize his own lifetime utility when he chooses a college.

## III. The College Admissions Project Survey Data

Our data comes from the College Admissions Project, in which we surveyed high school seniors applying to college during the 1999-2000 academic year.[3] The survey was designed to gather data on an unusual group of students: students with very high college aptitude who are likely to gain admission to and attract merit scholarships from selective colleges. While such students are represented in surveys that attempt to be nationally representative, such as the National Educational Longitudinal Survey, they are a very small share of the population of American students. As a result, the number of such students is always so small that their behavior cannot be analyzed, even if the nationally representative survey itself contains a large number of students. Yet, questions of the type that motivate this paper apply acutely to students of high college aptitude, who can–if they wish–consider a wide variety of colleges, merit scholarships, and aid packages. By focusing on students with very strong academic credentials, we can hope to learn how the students who can attract interesting packages of aid and scholarships respond to them.

### A. The Survey Design

In order to find students who were appropriate candidates for the survey, we worked with counselors from 510 high schools around the United States. The high schools that were selected had a record of sending several students to selective colleges each year, and they were identified using published sources (such as Peterson's guides to secondary schools), the college admissions experience of Andrew Fairbanks, and evidence from Avery's previous work. Each counselor selected ten students at random from the top of his senior class as measured by grade point average. Counselors at public schools selected students at random from the top 10% of the senior class, while counselors at private schools (which tend to be smaller and have higher means of measures of college aptitude) selected students at random from the top 20% of the senior class.[4] The counselors distributed the surveys to students, collected the completed surveys, and returned them to us for coding.[5] Students were tracked using a randomly assigned number; we never learned the names of the students who participated.

Survey participants completed two questionnaires over the course of the academic year. The first questionnaire was administered in January 2000. It asked for information on the student's background and college applications; the majority of these questions were taken directly from the Common Application, which is accepted by

---

[3] See Avery and Hoxby [2000] for additional detail.

[4] The counselors were given detailed instructions for random sampling from the top 20, 30, 40, or 50 students in the senior class depending on the size of the school. For example, a counselor from a public school with 157 students was asked to select 10 students at random from the top 20 students in the senior class, with the suggestion that the counselor select students ranked #1, 3, 5, 7, 9, 11, 13, 15, 17, and 19.

[5] The exception was the parent survey, which parents mailed directly to us in an addressed, postage-paid envelope so that they would not have to give possibly sensitive financial information to the high school counselor. Because counselors have access to the information on the students' surveys (and must, in order to support their applications competently), we were not as concerned about students' giving information to their counselors.

many colleges in place of application forms specific to those colleges. Each student listed up to ten colleges where he had applied, his test scores, and race. In addition, each student listed the colleges and graduate schools (if any) attended by each parent and the colleges (if any) attended by older siblings along with their expected graduation dates.

The second questionnaire was administered in May 2000 and asked for information about the student's admission outcomes, financial aid offers, scholarship offers, and matriculation decision. Each student listed their financial aid packages with the amounts offered in three categories: Grants, Loans, and Work Study. Each student also listed institutional scholarships (scholarships offered by a specific college for exclusive use there) and outside scholarships (and their restrictions, if any).[6] The response on merit-based scholarships, both institutional and outside, was particularly accurate and clear, presumably because students were proud of them as accomplishments. Finally, each student was asked an open-ended question: "Did finances play a role in your decision?"

A third questionnaire was distributed to a parent of each survey participant. The parent was asked to indicate whether either tuition or financial aid considerations (or both) would affect their child's choice of college. In addition, each parent was asked to check one of fifteen boxes to indicate their income range in 1999. (See Table 1 for the income categories.)

We matched the College Admissions Project data to administrative data on tuition, room, board, comprehensive cost, enrollment, and expenditure. In all cases, the ultimate source for the administrative data was the college itself and the data were for the 2000-01 school year, which corresponds to the survey participants' freshmen year.[7] Only for one variable did we encounter difficulty gathering complete administrative data: per-pupil expenditure disaggregated into instructional spending and educational spending. It is our hope that we will eventually obtain complete data for these variables by sufficient inquiry at the colleges themselves. However, we currently have expenditure data for only 72 percent of the matches between students and the colleges to which they have been admitted.[8]

The College Admissions Project survey produced a response rate of approximately 65%, including

---

[6] Students were offered the option of photocopying their financial aid offers, blacking out their names, and submitting the copy in place of answering the question. A small number of students did so.

[7] We collected the administrative data from the following sources in order: The College Board's annual survey, the United States Department of Education's Integrated Postsecondary Education Data System (IPEDS), the United States Department of Education's College Opportunities Online system (COOL), the 2001 edition of Peterson's Guide to Colleges, and colleges themselves. That is, we attempted to fill in each observation using the first source first; missing observations were filled in using one of the remaining sources, in order.

[8] We are able to remedy this data problem only with painstaking inquiry at individual colleges, but we do expect to have expenditure data on more than 95 percent of colleges eventually. We hope that readers will understand that the publication version of this paper will include more expenditure data, but that this version cannot.

information for 3,294 students from 396 high schools.[9] The final sample contains students from 43 states plus the District of Columbia.[10] Although the sample was constructed to include students from every region of the country, it is intentionally representative of applicants to highly selective colleges and therefore non-representative of American high school students as a whole. Regions and states that produce a disproportionate share of the students who apply to selective colleges are given a weight in the sample that is approximately proportionate to their weight at very selective colleges, not their weight in the population of American high school students. Of course, all of the students in the sample have very strong academic records. It is not surprising that the sample contains students whose parents have higher incomes and more education than typical American parents.

B. The Typical Student in the College Admissions Project

The summary statistics shown in Tables 1 and 2 (and in Appendix Tables 1 and 2) demonstrate that the sample is quite special. The average (combined verbal and math) SAT score among participants was 1357, which put the average student in the sample at the 90th percentile of all SAT takers. About 5 percent of the students won a National Merit Scholarship; 20 percent of them won an outside scholarship that was *fully* portable; and 46 percent of them won a scholarship from at least one college. 45 percent of the students attended private school, and their parents' income in 1999 had an average of approximately $119,929. However, 76 percent of the sample had incomes below the cut-off where a family is considered for aid by selective private colleges (the cut-off is approximately $160,000, but the actual cut-off depends on family circumstances). 59 percent of the students applied for need-based financial aid, and 41 percent of the families reported that finances influenced their college choice.[11] Of course, a college may offer a student a scholarship or grant to persuade him to matriculate, regardless of whether he has applied for aid.

83 percent of the student's parents were currently married, and 23 percent of the students had at least one sibling currently enrolled in college. The racial composition of the survey participants was 73 percent white non-Hispanic, 16 percent Asian, 3.5 percent black, and 3.8 percent Hispanic. We found that the black and Hispanic sub-groups were too small for substantial, separate analysis. We also found that the white and Asian sub-groups behaved similarly, all else equal. Thus, we will not discuss students' races further in this paper.

Looking at Appendix Table 1, which shows descriptive statistics on the colleges where the students *applied*,

---

[9] The most common reasons for failure to return the survey were changes of high school administration, an illness contracted by the counselor, and other administrative problems that were unrelated to the college admissions outcomes of students who had been selected to participate.

[10] The states missing from the sample are Alaska, Delaware, Iowa, Mississippi, North Dakota, South Dakota, and West Virginia.

[11] That is, either the parent, the student, or both claimed that finances influenced the college choice decision.

we can see that the survey participants applied to a range of colleges that included "safety schools" (the mean college to which a student applied had a median SAT score 8.5 percentiles below the student's own). However, the participants also made ambitious applications: 47.5 percent of them applied to at least one Ivy League college.

Table 2 shows descriptive statistics for colleges to which the students were *admitted*. This is the set of observations on which we concentrate in our analysis of college choice–for the simple reason that students can choose only among those colleges to which they were admitted. Comparing Table 2 to Appendix Table 1, we can see that the students made logical application decisions. The mean college to which they *applied* had a median SAT score at the 83rd percentile; the mean college to which they were *admitted* had median SAT score at the 81st percentile. This small difference suggests that the students aimed a little high in their applications, a procedure that is optimal. 66 percent of the colleges to which they were admitted were private, and their mean tuition was $17,671. Note that we summarize the colleges' in-state tuition, out-of-state tuition, and the tuition that actually applies to the students in the sample (in-state or out-of-state as appropriate).

Finally, Appendix Table 3 shows descriptive statistics on the colleges at which the students matriculated. They are more selective, on average, than the colleges to which the students were admitted: their median SAT score is at the 83.4th percentile, as opposed to the 81st percentile median SAT score of the colleges to which students were admitted. This makes sense because it implies that students included "safety schools" in their choice sets, but that they did not actually matriculate at their "safety schools" when they did not need to. One measure of the unusual college aptitude of the survey participants is the list of colleges at which the largest numbers of participants enrolled. Seventeen institutions enrolled at least 50 students from the sample: Harvard, Yale, University of Pennsylvania, Stanford, Brown, Cornell, University of Virginia, Columbia, University of California–Berkeley, Northwestern, Princeton, Duke, University of Illinois, New York University, University of Michigan, Dartmouth, and Georgetown.

C. Some Variables with Interesting Measurement Issues

Our measurement of most variables was perfectly straightforward, but a few exceptions are worth mentioning. We converted American College Test (ACT) scores to SAT scores using the cross-walk provided by The College Board. We converted all college admissions scores into national percentile scores using the national distribution of SAT scores for the freshman class of 2000-01.[12] We used longitude and latitude to compute the distance between a student's high school and each college to which he applied. We used parents' reports of their own incomes whenever available.

---

[12] This is an important, though often neglected, conversion. A given change in an SAT scale score (of, say, 100 points) corresponds to a differing number of percentiles depending where the scale score is in the distribution. For instance, the difference between a combined 1500 and 1600 is only a few percentiles, but the difference between a combined 1400 and 1500 is three to four times as many percentiles. Using the unconverted scale scores tends to wreak havoc with regression procedures in which the scale score enters linearly.

When a parent report of income was unavailable, we substituted an estimate of parents' income based on the Expected Family Contribution reported by the student. (The Expected Family Contribution is the standardized federal estimate of the amount that parents should be able to contribute towards the student's college education.) Parents' income and current expenditures for the college education of older siblings explain about 88 percent of the variation in the Expected Family Contribution. Therefore, our estimates of parents' income based on the Expected Family Contribution and siblings' college expenses (which we can estimate because we know about sibling enrollment) are highly accurate for our purposes. Readers will see that we only need to group parents into four coarse income groups. For those families that had both reported parents' income and an Expected Family Contribution, our estimate of parents' income based on Expected Family Contribution placed families into the correct group 97 percent of the time.

A remaining 3.4 percent of families had neither a reported parents' income nor a reported Expected Family Contribution. For these families, we estimated parents' income by assigning parents the mean incomes for people with the same detailed occupation in the March 2000 Current Population Survey (which asks about a person's income in1999, from his occupation). For families for which we could check this method, we found that it assigned them to the correct coarse income group 91 percent of the time.[13]

Finally, because the aid and scholarship variables are important, we hand-checked every observation to ensure that no scholarship was counted twice (as a grant and again as a scholarship), recorded incorrectly as the four-year total rather than the annual amount, or recorded with insufficient restrictions. In many cases, we were able to double-check or clarify students' responses because they were offered named scholarships with known parameters (for instance, "Morehouse Scholars" at the University of North Carolina).

## IV. Empirical Strategy

Our empirical strategy is straightforward. We are interested in discovering the factors that influence a student's choice among the colleges to which he is admitted. This is an estimation problem for which *conditional logit* (also known as McFadden's Choice Model) is ideally suited.

Intuitively, conditional logit groups together the colleges to which each student was admitted. This becomes a student's menu or college choice set. A binary outcome variable shows which college was actually

---

[13] For the occupation-based estimate of parents' income, nearly all of the mistakes would have been caused by our assigning families to the medium-high income category when they truly belonged in the high income category. (We are referring to the coarse income categories defined later in the paper.) Thus, we suspect that our medium-high income group probably contains about 12 families that should be assigned to the high income group. The under-assignment to the high income group is caused by professional occupations', such as being an attorney, having income distributions with a right-hand skew. Thus, parents make it into the high income category because they are, say, unusually highly paid lawyers, not because the mean income for a lawyer would put them there.

picked–in our case, it is a dummy variable equal to 1 for the college at the student matriculated and 0 for all of the other colleges in his choice set. Each college in the choice set has a number of attributes, some of which are the same for all students (such as whether the college is public or private) and some of which depend on the identity of the student (such as grants to the student from the college). Conditional logit estimation relates the binary outcome variable to the college attributes by maximizing the following log likelihood function:

$$\ln L = \sum_{i=1}^{n} \sum_{j=1}^{J_i} matric_{ij} \ln Prob(collegechoice_i = j) \ , \tag{3}$$

where

$$Prob(collegechoice_i = j) = \frac{e^{\beta' x_{ij}}}{\sum_{j=1}^{J_i} e^{\beta' x_{ij}}} \ . \tag{4}$$

Examining equation (3), one sees that the conditional logit estimates are those that maximize the *similarity* of the estimated likelihoods and the actual matriculation decisions.

In equations (3) and (4), *i* indexes the student; *j* indexes the college; the indicator variable $matric_{ij}$ is equal to one if student *i* chooses to matriculate at college *j*, and zero otherwise; and $collegechoice_i$ is simply the student's college choice. The vector $x_{ij}$ includes the attributes of choice *j* for student *i*–note that the subscripts indicate that the attributes may be match specific. $\beta$ is the vector of effects that we are interested in estimating. One maximizes the *log* of the likelihood simply to make estimation easier.

The choice problem we are investigating is ideal for conditional logit estimation but unsuitable for multinomial logit or probit estimation. Although multinomial logit and probit are related to conditional logit and sometimes confused with it, they cannot be used to examine choice in a situation where choices have match-specific attributes, such as a scholarship that applies to one student accepted by a college but not to all students accepted by that college.

There are a number of empirical issues that deserve comment.

A. The Variation that Drives the Estimates

It is important to be explicit about the variation that drives our estimates. There is variation in the attributes of colleges in students' choice sets mainly because colleges vary for reasons for their own that are effectively exogenous to the individual student. For instance, colleges differ in location, in whether they are publicly or privately controlled, in endowment, in the niche they fill in the market for college education, and so on. All of this and much more variation in their attributes is effectively exogenous or parametric to the individual student, who must accept the range of choices that exist.

There is one possible worry about the endogeneity of the attributes of college choices, but it seems minor based on *a priori* grounds as well as empirical evidence. It is as follows. We might worry that a college's aid offer to a student is not exogenous because it reflects effort on the student's part *that is observed by that college only* and

that is a function of how much his desire to attend that college. For instance, a student might get aid offers from several colleges and find that his most preferred college is offering slightly less aid than his second most preferred college. He might send special letters or other fresh evidence of merit to his most preferred college (and only his most preferred college) in order to induce them to give him a slightly better aid package. Thus, the aid we ultimately record for his most preferred college might be a function of a match-specific liking that we do not observe. There are three reasons to think that this form of endogeneity is minor. First, colleges decide on a student's aid package in a calculating way, taking his merit, his need, and the college's own objectives into account. Although a student who makes extra effort may improve his aid package somewhat, he is unlikely to get it changed radically because he has already revealed most of his achievements in the applications he sent to all colleges. In other words, most of a student's merit is already fixed and revealed by the time the student can begin an effort campaign directed at a single college. Second, in our survey data, there is little indication that students were able to make special efforts that convinced colleges to give them substantially different aid offers. The College Admission Project survey asked students whether they had been able to get any aid offer revised. Revisions occurred in only 9 percent of possible college choices, and most of the revisions were minor in character, based on the verbatim responses of students who described the revision they obtained. Only 2 out of the 3,240 students described a revision that would cost a college more than $1,000.

B. The Role of a Student's Own Attributes

People are sometimes surprised to find that there are no individual student attributes (such as the student's SAT score) included in the vector $\boldsymbol{x}_{ij}$ for conditional logit estimation. However, a little thought shows why this is so. The student's own characteristics are same regardless of the choice he makes, so they cannot be a reason for choosing one college over another. It is only college attributes or *match-specific* attributes (think of interactions between a student's characteristics and a college's attributes) that can influence his choice. For instance, a student might care about whether his SAT score is much higher or lower than the average SAT score at a college. Thus, measures of the difference between a student's SAT score and a college's average SAT score are included in the vector $\boldsymbol{x}_{ij}$. Also, college attributes that differ across colleges but are constant across students within a college (such as whether the college is publicly or privately controlled) may be included in the vector $\boldsymbol{x}_{ij}$ since they obviously can be reasons for choosing one college over another.

Individual student attributes may affect college choice even though they are constant across all of a student's choices. This is because they may affect the *way* that he responds to a particular college or match-specific attribute. For instance, a student from a low income family may be more responsive to loans offered by a college than is a student from a high income family (which presumably has many more alternatives to the college's loan, including regular bank or home equity loans). Thus, although we cannot include student attributes as $\boldsymbol{x}_{ij}$ variables that affect

college choice, we do estimate our choice model separately for students with different attributes. For instance, we show tests for different responsiveness of students with different family income, of students whose parents attended more and less selective colleges, and of students from private and public secondary schools. In fact, we tested for different responsiveness along many other dimensions; if we do not show a dimension it is simply because there was no sign of differential responsiveness.

C. The Consequences of Observing Only a Subset of a Student's Possible College Choices

We do not observe all of the colleges to which a student *could* have been admitted and all of the financial aid packages they would have offered him. We focus on the subset of colleges to which he was admitted, among those to which he applied. One might think that some of bias would arise from this limitation, but no bias does. Indeed, the estimates are far *more* likely to be consistent than they would be if we were to include all colleges in students' choice sets (not just those to which they were admitted). Formally, this is because the independence of irrelevant alternatives restriction (a well-known implication of the conditional logit specification) is more likely to be satisfied. Within the conditional logit literature, our application is almost ideal: the data and facts match the assumptions in the econometric model to an almost unique degree. In the Econometric Appendix, we offer more detail on the econometric issues.[14] Here, we offer a purely intuitive version.

Suppose that we *did* observe all of the colleges to which student could have been admitted and all of the financial aid packages he could have received at each college, with a probability attached to each financial aid package.[15] Then, we would quickly find that, given the student's preferences, some colleges were dominated by others (or dominated except under extremely unlikely scenarios). The dominated colleges would be irrelevant to the student's choice. This is precisely what a student is revealing when he does not apply to a college: the college is dominated. Note that a college could be dominated simply because it was very similar to (but very slightly less preferred than) other colleges in the student's choice set. For instance, many students apply to a "safety school" to which they know they will be admitted with near certainty. However, no student applies to numerous, very similar "safety schools." Some of the "safety schools" would be dominated and therefore irrelevant to his choice. When a student chooses to apply to a school, he is essentially stating that the college could be a relevant alternative. By forming his choice set this way (called "endogenous choice set formation" in the econometric literature), he is helping to guarantee that the restrictions for conditional logit are actually satisfied (specifically, the independence of irrelevant alternatives is satisfied). For obtaining consistent estimates, it is *ideal* to work with a person's endogenous

---

[14] Specifically, we address three issues: the independence of irrelevant alternatives, endogenous choice sets, and lack of balance in choice sets that arises naturally when choice sets are endogenous.

[15] For instance, the student might have only a 10 percent probability of receiving a special scholarship of $10,000 at college *j*, but an 80 percent probability of receiving a small grant of $1,000 there.

choice set. We are fortunate to be in this position.

To summarize, we do not know *why* some colleges are dominated and thus irrelevant, but students' application behavior eliminates irrelevant colleges for us. This is not only acceptable, but ideal for consistent estimation. This is because the conditional logit model imposes independence of irrelevant alternatives, and the best way to satisfy this restriction is to exclude irrelevant alternatives from the choice set.

D. Early Decision Applicants

While we generally benefit by looking at choice sets composed exclusively of colleges to which a student was admitted, we are deprived of information when a student applies through an early decision program and is accepted. In such cases, we the econometricians arrive too late on the scene. Essentially, the student has decided to make his college choice *prior* to the formal application process, in return for the admission benefits of an early decision application and the option of applying elsewhere should he be rejected in the early round. Because some early applicants apply to and are accepted by only one college (being accepted in December by their most preferred college, they have no reason to file regular applications due in January), we sometimes do not know which college alternatives the student considered to be relevant when he chose the college at which he ultimately matriculated. In other words, we lose the opportunity to study the student's choice decision, just as if we had accidentally lost the data on him. Indeed, readers may find it helpful to think of the one-acceptance, early application students as students with lost data. Note, however, that a good many early decision students *do* indicate the relevant alternatives by applying to them through the regular application process before they learn about their early decision outcome. Such students' college choices are observable to us, and we use them. Out of 3,240 students, 338 (or 10.4%) are early applicants for whom we do not observe other regular college applications. This group includes a few early *action* (not early decision) applicants who are not constrained to matriculate at the college that accepts them in December.

Our basic set of estimates does not use the behavior of these 338 early applicants who apply to only one college. (The estimation procedure automatically sets them aside because there is no variation in the outcome $matric_{ij}$; we do not have to purposely exclude them.) However, excluding the 338 students is undesirable because excluding them is not like excluding students at random–they might be an unusually risk-averse or sophisticated group of students.

We attempt to remedy the problem by showing a second set of results based not the actual colleges in a student's choice set but in an extended choice set that also includes other colleges *predicted* to have been in his more fundamental choice set. Using predicted choice sets is the procedure recommended in the econometric literature on conditional logit estimation with endogenous choice sets. We have unusually good conditions for forming predicted choice sets because we do observe choice sets containing multiple colleges for the vast majority of our survey respondents. Moreover, we always observe choice sets for similarly qualified students *from the same school.* This

fact is a fortuitous consequence of our sample design, but it is very useful because a primary problem in predicting endogenous choice sets is getting *local* information right. For instance, a student who applies to Princeton, Stanford, and Johns Hopkins from Birmingham, Alabama is likely to include locally salient alternatives from the southeastern United States in his choice set, while a student who applies to the same three colleges from Eugene, Oregon is likely to include locally salient alternatives from the northwestern United States. Predicting which local colleges are considered *relevant* by high achieving students in a school might be difficult, but we are able to observe the local alternatives that tend to be picked.

Specifically, we use the following procedure to form the predicted choice sets. We preliminarily assign to each student a choice set equal to his own observed choice set, plus all other colleges to which at least one of his schoolmates applied. Call these extra colleges the "add-on colleges" for convenience. We then need to eliminate, from the student's add-on colleges, colleges that are unlikely to admit him. We do this by eliminating all add-on colleges where the student's own SAT score would put him below college's median SAT score. (We experimented with other percentile thresholds down to the 25th percentile, but we found that the results did not change much.) We now need to create reasonable aid packages for a student's add-on colleges. Because aid tends to have a local character (a highly meritorious student from Detroit will systemically receive a different aid package at the University of Michigan than a highly meritorious student from Kansas City), we create an aid packages for each student at each add-on college by using the aid package of another students in his school who *was* admitted to that colleges. Because grants depend mainly on merit, each add-on college's grant is filled by the grant actually received by the other student who (1) came from the student's own high school , (2) was accepted by the college, and (3) had the SAT score most similar to the student's own. Because loans and work-study depend mainly on parental income, each add-on college's loans and work-study are filled by the loans and work-study actually received by the other student who (1) came from the student's own high school , (2) was accepted by the college, and (3) had parental income most similar to the student's own.

We show the results based on these predicted choice sets, but only after presenting our basic results. So long as the two sets of results are similar, we can be reasonably confident that our evidence does not hinge on the exclusion of early applicants accepted by only one college. We are interested in whether the results are similar, not identical (or nearly identical). We do not expect the results to be identical because the estimates from the predicted choice sets are likely to be slightly inconsistent for two reasons: they are based on explanatory variables that are measured with error for the add-on colleges (attenuation bias), and they are more likely to violate the independence of irrelevant alternatives assumption.

E. The Specification of $\beta' x_{ij}$

We have not attempted to restrict the set of variables that affect college choice in any way. We have

included all variables available to us that seemed at all likely to affect students' choices. We are willing to include any other variables that might be suggested to us, subject only to data availability.

We have imposed only two restrictions that seem worth mentioning. First, we measure all the financial variables in thousands of annual dollars, not in the natural log of dollars or other transformation. This is because we wish explicitly to test whether students react similarly to the same dollar amount when it arises in two different, but fundamentally similar, forms. For instance, a rational student following the classic human capital investment model would be expected to react similarly to a reduction of $1,000 in the annual tuition of a college and an increase of $1,000 in the annual grant given him by the college.[16] Also, we wish explicitly to test whether students react differently to the same dollar amount when it arises in two forms that cost the college very different amounts. For instance, the aforementioned rational student should not react similarly to $1,000 in grants and $1,000 in loans. The cost to the college of a loan may be anywhere from zero to about 15 percent of its face value, but it is rarely if ever close to 100 percent of its face value (as the grant is).

Second, we largely restrict the $x_{ij}$ variables to having a linear effect. We have made this restriction purely because the greater simplicity facilitates interpretation. We have allowed some variables, such as distance and a student's SAT "match" with a college, to have nonlinear effects. We would be willing to consider allowing others to have nonlinear effects also, especially if a particular non-linear specification is suggested to us.

F. The Interpretation of the Estimates We Show

We display the conditional logit results using odds ratios and *Z* statistics. An odds ratio gives us the ratio of the posterior odds of a college choice to the prior odds of a college choice *when only the variable in question is allowed to change*. For instance, we could compute the odds that college *j* is chosen; then raise its tuition by $1,000, holding all other variables equal, and recompute the odds. The former odds would effectively be the prior odds, and the latter odds would effectively be the posterior odds. In short, the odds ratio is $\alpha$ in the following expression:

$$\text{posterior odds} = \alpha \cdot \text{prior odds}.$$

It is convenient to compute the odds ratio from our estimated vector $\hat{\beta}$ because the odds ratio is just $e^{\hat{\beta}}$.

The way to interpret the odds ratio for a certain variable in $x_{ij}$ is the proportional change in the odds of student *i* attending college *j* for a unit increase in the variable, holding all other variables constant. A simple example would be the indicator for whether a college is public. If the odds ratio is 1.1, then a one unit change in the variable (corresponding to a switch from private to public) would make the student's odds of attending the college 1.1 times whatever the odds were if the college had been private. Note that the change is proportional to the prior

[16] In this case, the reactions need not be identical because a reduction in tuition may affect more and different *other* students than grants do. Thus, the student's beliefs about the college's revenues may depend on whether it is reducing tuition by $1,000 or increasing grants by $1,000. Of course, these two actions do have the same direct effect on a student's college cost.

odds: if the student's prior odds of attending the college were 30 percent, the posterior odds would be 33 percent (30 times 1.1). If the student's prior odds of attending the college were 80 percent, the posterior odds would be 88 percent (80 times 1.1).

Naturally, an odds ratio greater than one means that an increase in the variable *raises* a student's probability of matriculating, all else equal; an odds ratio less than one means that an increase in the variable *reduces* a student's probability of matriculating, all else equal.

The *Z* statistics are easy to interpret. They are akin to the more familiar *t* statistics and have the same thresholds for statistical significance. Thus, an odds ratio with a *Z* statistic greater than 1.96 is statistically significantly different from 1 (the null hypothesis of no effect) with 95 percent confidence, and so on.

## V. How Aid Affects College Choice

### (Our Basic Results on the Determinants of College Choice)

In this section, we discuss our basic results on the determinants of college choice, which are presented in Table 3. Recall that the outcome is matriculation, a binary variable equal to 1 for exactly one of the colleges to which a student was admitted.

If we examine Table 3's overall pattern of signs and statistical significance, students' college choice appears to be very reasonable. Students are more likely to attend a college if, all else equal, it offers them larger grants, offers them larger loans, offers them a larger amount of work study, is the most selective college to which they were admitted, is their father's *alma mater*, or is the same college that their sibling attended or attends. Students are less likely to attend a college if, all else equal, its tuition is higher, its room and board is higher, its mean SAT score is more below theirs, or it is the least selective college to which they were admitted. Several variables do not have a statistically significant effect on students in our survey: the amount by which the college's average SAT exceeds the student's, an indicator for the college being their mother's *alma mater*, the distance between the college and the student's high school, whether the college is public, and whether the college is in-state.

At the broad "sign and statistical significance" level, the results seem to us to be very close to our expectations. Are they so close to our expectations, however, when we examine the odds ratios in detail?

The left-hand column of Table 3 shows us that an additional thousand dollars in grants raises a student's probability of matriculating by 11.0 percent of his prior probability; an extra thousand dollars in tuition lowers a student's probability of matriculating by 2 percent of his prior probability, and an extra thousand dollars in room and board lowers a student's probability of matriculating by 9.1 percent of his prior probability. While these results hint at students' being more sensitive to aid than to reduced tuition, we cannot reject the hypothesis that students respond equally to the same dollar amount in the form of increased grants, reduced tuition, or reduced room and board.

What is more surprising is the response to loans and work study. An additional thousand dollars in loans raises a student's probability of matriculating by 9.8 percent of his prior probability, and an additional thousand dollars in work study raises a student's probability of matriculating by 41.1 percent of his prior probability. (Note that the work-study response is imprecise and is therefore not statistically significantly different from the response to grants.) These results suggest that students do not view loans and work study as much inferior to grants, despite the fact that they cost a college much less than a grant does. Loans and work study have substantial costs for students, even though the burden of loans is delayed and the cost of work study is in the form of foregone leisure.

For every percentile that a student's SAT score *exceeds* the mean SAT score of a college, his probability of matriculating falls by 4.9 percent of his prior probability. Although the odds ratio on the percentile that a student's SAT score *fall short of* the mean SAT score of a college is not statistically different from one, its point estimate is greater than one, suggesting that students are not deterred by a college's having higher average SAT scores than their own. That is, the SAT match variables are not symmetric, but indicate that students only dislike a mismatch if their own SAT scores are "too high" for the college. Along similar lines, a student's probability of matriculating rises by 76.2 percent of the prior probability if the college is the most selective among the colleges to which he was admitted, and it falls by 29.9 percent of the prior probability if the college is the least selective among the colleges to which he was admitted. Overall, these results strongly suggest that students place substantial weight on a college's selectivity as a measure of its value. This is not necessarily because students value selectivity *per se* rather than a college's resources; it may simply be that selectivity is highly correlated with resources and that selectivity is easier to observe (and probably better measured) than resources.

Finally, a father or sibling who attended the college greatly increases a student's own probability of attending it: having a father or sibling attend raises the probability of matriculating by, respectively, 71.5 percent and 94.9 of the prior probability. This strong family alumnus effect may be due to the student's familiarity with or allegiance to the college, but it might equally be match-specific attributes that are similar for the student and other members of his family (shared tastes, similar career plans, and so on).

The right-hand column in Table 3 adds an additional variable, per-pupil instructional spending, that should affect student's college choice given that tuition and aid have already been shown to affect his choice. Per pupil instructional spending is our best measure of the actual dollar value of the resources invested in the student's human capital. This is particularly true for the students in this sample because the public flagship universities to which they apply have per-pupil spending that consistently exceeds in-state tuition and the private colleges to which they apply have tuition that is subsidized by endowment funds. The degree to which tuition is subsidized varies greatly, at both public and private colleges. Thus, per-pupil instructional spending should be an important factor that picks up resources not directly measured by any other variable in the regression.

We can see that per-pupil instructional spending appears to have little or no effect on college choice. The estimate is far from being statistically significantly different from 1, and the point estimate of the odds ratio is 0.999, very close to 1. What are the possible explanations for instructional spending's lack of influence? It is unlikely that students do not actually care about the instructional resources they experience, given the time they spend working with these resources and the pressure they put on college administrations when the resources do not match their expectations. It is also unlikely that instructional spending is disbursed so inefficiently at high spending colleges and so efficiently at low spending colleges that they really offer similar resources. Instructional spending varies *greatly* among colleges: from $2,000 to $51,000, with a standard deviation of $10,561 (see Table 2). It is implausible that high spending colleges are so inefficient that their resources are equivalent to those of colleges that spend less than one twenty-fifth of what they spend.

What are the likely explanations for instructional spending's lack of influence? There are two, and each one probably explains part (they are *not* mutually exclusive). First, instructional spending is imprecisely measured because it is not measured separately for undergraduates. (A university's instructional spending is put onto a per-pupil basis by dividing it by full-time equivalent students, undergraduate and graduate.) Imprecise measurement of instructional spending would produce attenuation bias –that is, the odds ratio being biased toward 1. The odds ratio shown is certainly consistent with attenuation bias.

Second, we know that college selectivity is highly correlated with measured instructional spending, and college selectivity may be *more* highly correlated with a college's real resources than is the measure of per-pupil instructional spending. This might occur if students were drawn by the true measure of available resources. Under this scenario, the odd ratios associated with the four selectivity variables represent the *combined* draw of higher aptitude students and greater resources.

Note that the right-hand column of Table 3 uses only 5513 (or 60.5 percent) of the 9113 observations used in the the left-hand column. This is the unfortunate consequence of the poor availability of the instructional spending variable (which we hope to remedy in future versions of this paper). However, the smaller number of observations in the right-hand column is probably not the cause of instructional spending's being statistically insignificant. Most other variables in the regression retain their odds ratios and their statistical significance in the smaller sample. Indeed, a regression identical to that in the left-hand column was estimated using just the smaller sample, and the results are very similar to those for the larger sample. For conciseness, these extra estimates are not shown.

## VI. Are Early Decision Students Different?

In Table 4, we compare our basic results to results that include early applicants accepted by only one college. Specifically, the left-hand column of Table 4 repeats our basic results from the left-hand column of Table 3,

and the right-hand column includes the early applicants accepted by only one college, substituting their predicted choice sets for their actual choice sets. We described the procedure for generating predicted choice sets in Section IV.

The table shows that the results change *very* little when the early decision students are included, with their predicted choice sets. The only coefficient that changes non-negligibly is the coefficient on the college's being the least selective college in the choice set. In the basic results, this coefficient indicates that the least selective college is *not* attractive; in the results that include predicted choice sets, this coefficient indicates that the least selective college *is* attractive. This change in the coefficient is certainly due to our mistakenly predicting that a student could be admitted to colleges where he would be rejected. When we make such mistakes in our predictions, a student will occasionally appear to attend the least selective college in his choice set–in fact, the other colleges in his predicted choice set are probably not really available choices. Put another way, we violate the independence of irrelevant alternatives restriction when we add mistakenly colleges that are not actually in a student's choice set; we should expect some inconsistent estimates.

Because including the early applicants who have only one college in their choice set makes so little difference to our results, we do not show results for predicted choice sets from here onwards.[17]

## VII. Do Students from Different Families Make College Choices Differently?

In this section, we investigate whether students from different family backgrounds make college choices differently. That is, do they respond differently to the same college and match-specific attributes? By design, the students in the College Admissions Project sample are similar in aptitude, but their background are much less similar along the dimensions of family income, parents' college experience, and so on. In Tables 5 through 8, we show the results of reestimating our basic conditional logit specification (the left-hand column of Table 3) for different subsets of students. We do not include per-pupil instructional spending in these regressions because we found that it was never statistically significant, and including it substantially reduces the sample.

### A. Students with Different Family Income

One obvious hypothesis is that students with different family incomes will respond differently to aid, tuition, room and board, and other attributes of colleges. Because high income families can more easily finance college out of savings or obtain cheap loans, we suspect that students from such families will be less sensitive to the

[17] Readers may be interested to know that if we use predicted choice sets for *all* students (not just early applicants), we obtain clear evidence of the mildly inconsistent estimates that we expected. Specifically, the coefficients on the aid variables show evidence of slight attenuation bias, which we expect because they are only estimates for many observations. We also see more evidence of violations of the independence of irrelevant alternatives, such as coefficients on the distance variables that suggest that irrelevant local colleges are too often being included in students' predicted choice sets.

variables that determine how much they will actually pay for college in any given year.

For Table 5, we divided students into four groups based on family income: "low" being less than $40,000; "medium low" being $40,000 to $80,000; "medium high" being $80,000 to $140,000; and "high" being greater than or equal to $140,000. The right hand column in the table contains the word "rejected" when the hypothesis that the odds ratios for the four income groups are equal is rejected with 95 percent confidence.

Students from families with low to medium high incomes respond to $1,000 in grants by raising their probability of matriculation by about 13 percent of their prior probability. Students from high income family also respond, but only by raising the probability of matriculation by 7.5 percent of their prior probability. We can reject the hypothesis that the 13 percent and 7.5 percent responses are the same, with 95 percent confidence.

We cannot reject the hypothesis that income groups respond identically to $1,000 of loans, although the point estimates hint that students from low income families are more sensitive. While we can reject the hypothesis that $1,000 in work-study has the same effect on students from all income groups, the pattern of estimates is difficult to interpret. Students with low and medium high family incomes appear to be less sensitive than students from medium low and high family incomes. Indeed, we generally find our estimates of the effect of work-study difficult to interpret in this paper. We suspect that this may be because work-study has extremely important attributes, contained in the job itself, that we do not observe. For instance, it may be that work-study is an academically valuable research job at some colleges, a light-duty desk job at another, and an onerous cleaning job at yet another.

We find that students from medium low and medium high income families are more sensitive to tuition than students from families with low or high incomes. This makes sense because we expect middle income parents to have a more difficult time paying tuition than do high income parents. Low income parents are likely to rely on need-based aid and rarely pay the marginal tuition dollar anyway; they may therefore be quite insensitive to tuition differences.

As family income rises, students become less willing to attend a college where their own SAT scores exceed the college's mean score. The draw of the most selective college in a student's choice set rises with his family income; the repulsion of the least selective college in the choice set also rises with family income. For instance, a student in a high income family has his probability of matriculation rise by 109.3 percent of his prior probability for the college that is the most selective in this choice set. The corresponding number for a student from a low income family is 52.5 percent –less than half the reaction.

Compared to other students, students from high income families respond less to having a father or sibling who attended the college. Only low income students respond to a college's distance from their secondary school, and only medium high income students respond to a college's being in-state.

B. Students whose Parents Attended More and Less Selective Colleges

It may be that college choice differs between families with more and less experience of selective colleges. We test this hypothesis in Table 6, where parents are divided into groups based on the selectivity of their colleges (the maximum of the two colleges' selectivity).[18] Because the students themselves are solid applicants for *very* selective colleges, we were particularly interested in parents' experience with such colleges. Therefore, our high selectivity group contains parents whose college has a median SAT score at or above the 90th percentile; our medium selectivity group contains parents whose college has a median SAT score at or above the 70th percentile and below the 90th percentile; and our low selectivity group contains all other parents.[19]

We find that students whose parents attended low selectivity colleges are more responsive to grants and loans than students whose parents attended high selectivity colleges. For instance, $1,000 in grants raises the probability of matriculation by 12.3 percent of the prior priority for a student with "low selectivity parents" but the corresponding number is only 8.0 percent for a student with "high selectivity" parents. It appears that students with high selectivity parents are much more responsive to work study than other students, but we hesitate to give this result a firm interpretation (much as we hesitate over our other work study results).

The difference in response to tuition is impressive: while students with low and medium selectivity parents are repelled by higher tuition (each additional $1,000 in tuition reduces their probability of matriculating by 2.8 to 4.7 percent of their prior probability), higher tuition appears to attract students with high selectivity parents (each additional $1,000 in tuition *raises* their probability of matriculating by 4.2 percent of their prior probability). It may be that tuition appears to be attractive because it is correlated with per-pupil expenditure, a variable that should ideally be in the regression, as previously noted. In any case, it is clear that parents from colleges of higher selectivity have a less negative response to increased tuition.

We also find that, compared to other students, students with high selectivity parents are less attracted to their sibling's college, less attracted to in-state colleges, and are more attracted by a college's having a median SAT score that exceeds their own.

---

[18] There appear to be no students in the sample from families in which *neither* parent attended college. There are, however, a good many parents who have degrees from institutions that grant only the associate's degree or another degree below the baccalaureate.

[19] We were able to include Canadian colleges in the medium and high selectivity groups by approximating their selectivity. However, all other international colleges were included in the low selectivity group, in part because we are trying to measure parents' experience with selective *American* colleges, which admit students, grant aid, and charge tuition in manner that differs widely from other colleges around the world, including some colleges that are very selective in their own country. Only 1.1 percent of families have two parents who attended a non-Canadian international college.

C. Students from Public and Private Secondary Schools

In Table 7, we investigate whether students from public and private high schools make college choices differently. We find that an extra $1,000 in room and board reduces the probability of matriculating by 17.6 percent of prior probability among public school students, but that it has no discernable effect on private school students. We find that private school students are more repelled than are public school students by a college's being the least selective that admitted them. For public school students, being the least selective college in the choice set reduces the probability of matriculation by 20.4 percent of the prior probability. For private school students, the corresponding number is a much larger 44.2 percent. Being an in-state college is an attraction for public school students, whose probability of matriculation rises by 29.8 percent of their prior probability. Being in-state has no such draw for private school students (the point estimate indicates that they are slightly repelled, but it is not statistically significantly different from one). Interestingly enough, a college's being public attracts private school students but has no such effect on public school students.

D. Summing Up the Differences in College Choice among Students From Different Backgrounds

Overall, we find that students from high income families, whose parents attended more selective colleges, and who themselves attended private high schools are less deterred by college costs and less attracted by aid. They are also more attracted by a college's being selective, either because they are more attracted by the resources correlated with selectivity or because they are more attracted by high aptitude peers. We might ask, however, whether the differences in responsiveness shown in Tables 5 through 7 really add up to much. One way to answer to answer this question is to investigate whether students would alter their college choices if we made them act in accordance with the estimated model for *another* group of students. To create Table 8, we use the low income students' coefficients with the high income students' data, and *vice versa*. We do same exchange for students with low and high selectivity parents and for students from private and public schools. We show the percentage of students who would be predicted to change their college choice if they were to act in accordance with another group's model.

We find that a sizable minority of students *would* choose a different college within their choice set if they were to act like another type of student. Most notable are high income students, 41.1 percent of whom would matriculate at a different college if they made choices the way that low income students do. Similarly, 32.0 percent of low income students would alter their college choice if they behaved as high income students do. The corresponding numbers are in the range of 15 to 17 percent when we "exchange" the models of public and private high school students and are in the range of 21 to 23 percent when we exchange the models of students whose

parents attended low and high selectivity colleges.[20]

We should report, however, that while students' choice behavior is affected by variables like parents' income, parents' college selectivity, and private high school attendance, many other background differences do *not* appear to affect students' college choices. We tried and failed to find significant differences in choice behavior along dimensions like region of the country, recipiency of an outside scholarship like the National Merit, a record of leadership while in high school, and so on. We suspect that students differ *systemically* in college choice behavior mainly when they face constraints that are not easily overcome–parents' income and parents' willingness to pay for private education (which is probably correlated with parents' own college selectivity and willingness to pay for private high school).

## VIII. Do Students Respond to Aid Variables They Should Ignore?

So far, we have had only one test of whether students are responding to aid variables as models of human capital investment suggest that they should: our test was whether students responded as differently to grants, loans, and work study as they should, given the very different degree of subsidy incorporated in these three forms of aid. Students failed this test: they responded nearly identically to every additional $1,000, regardless of whether it was a grant or loan. The point estimates indicate that they were *more* attracted to $1,000 of work-study than $1,000 of grants, though this difference is not statistically significant. Although we have already mentioned that it is hard to interpret the coefficients on work-study without knowing what the jobs are, it is clear that work-study always requires some sacrifice of leisure, whereas grants require none. At this point in the paper, we have not made calculations that allow us to judge whether students respond too much or too little to grants, but *given* their response to grants, their responses to loans and work-study are certainly too large.

Furthermore, Tables 5 through 7 show that, while some students (high income, high parents' college

---

[20] We might also ask *how* the characteristics of the students' colleges would change if they were to alter their choices. That is, would students merely choose another college that was indistinguishable from their initial college? It is not possible to answer this question satisfactorily given the limited exercise we are attempting at this point in the paper. The reason we cannot give a satisfactory answer is that we are constraining students to re-choose *within* the set of colleges to which they applied and were admitted. Given our current purpose, which is merely to give readers a sense of the scale of the estimates in Tables 4 through 6, this constraint is temporarily acceptable. Later, when we attempt more ambitious thought experiments, it will be important to relax this constraint. Students would apply to a different range of colleges if they were to foresee themselves acting differently when it came to choosing a college. For example, if a low income student were to foresee that he would act like a high income student when he chose a college, he might apply to some high tuition colleges that he currently omits.

We can look at two outcomes in a satisfactory maner, however, because they are relative ones: the share of student who matriculate at the most and least selective colleges within their choice set. Here, we find that the altered choices are highly distinguishable. For instance, 94.7 percent of high income students choose the most selective college in their choice set when they act like themselves, but only 57.4 percent would do so if they were to act like low income students. 18.2 percent of low income students choose the least selective college in their choice set when they act like themselves, but only 6.8 percent would do so if they were to act like high income students. And so on.

selectivity, private high school) are less sensitive to aid, *all* students respond too much to loans and work-study, in comparison to grants. For instance, high income students treat grants, loans, and work-study at least as equally as low income students do.

A. Aid Variables that Students "Should Ignore"

In this section, we investigate whether students respond to aid variables they should largely, if not completely, ignore. Specifically, we look at three aspects of a grant. The first is whether the grant is called a scholarship. In other words, did the student merely report $4,000 in grants, or did he also report that the $4,000 was the, say, "Jane Doe Scholarship for Merit"? Based on the survey data, it appears that some colleges systemically name their grants, while others give similar amounts with no name attached. In fact, the correlation between the amount of a grant and its being called a scholarship is negative: -0.206.

We also examine whether a grant is front-loaded, so that the student receives more in his freshman year than in later years. The students in our survey, all of whom have an extremely high probability of completing college, should ignore such front-loading and simply look at the total amount of grants. We characterize a grant as front-loaded if the (nominal dollar) amount for the freshman year is greater than the (nominal dollar) amount for later years.[21] As a rule, front-loaded grants are heavily weighted toward the freshman year (for instance, $10,000 for one year, and $2,000 thereafter), not steady declining over the college career (for instance, $5,500, $4,500, $3,500, and $2,500 for the four successive years). Keep in mind that we will be investigating the effect of front-loading, holding the amount of the grant constant. Thus, we ask whether students respond more to, say, $4,000 per year if it is front-loaded. As a matter of fact, the correlation between the amount of a grant and its being front-loaded is negative: -0.189.

The final aspects of grants that we examine is what *percentage* they are of tuition, comprehensive cost, and per-pupil instructional spending. Obviously, if all colleges had similar tuition, comprehensive cost, and instructional spending, larger grants would always represent a higher percentage of costs and expenditures. But, colleges in our sample vary dramatically in tuition, comprehensive cost, and instructional spending. The standard deviation of in-state tuition is $9,594; the standard deviation of in-state comprehensive cost is $10,368; and the standard deviation of instructional spending is $10,561. Students in the sample applied to 40 colleges at which per-pupil instructional

[21] We do not assume that students should understand discounting. That is, we do not characterize a grant as front-loaded if it offers the same nominal amount for each year of college, even though the present discounted value of the freshman year grant is greatest. Also, we do not characterize a grant as front-loaded if it merely has conditions for continued good performance. For instance, a grant might be $4,000 for the freshman year, which will be renewed thereafter so long as the student maintains a B average. Such conditions are clearly intended to maintain achievement, rather than exploit students' impatience or myopia. Many graduate and professional programs use front-loaded grants because students' ability to earn money or win outside grants rises steeply during a student's graduate career–think of law schools, business schools, or Ph.D. programs. The same considerations do not apply to undergraduate programs.

spending was less than $3,500, and they applied to an equal number of colleges at which per-pupil instructional spending was more than $14,000, which is *four times $3,500*. Students in the sample applied to 106 colleges that have in-state tuition below $2,500 and applied to 143 colleges that have in-state tuition above $20,000, which is *eight times $2,500*.[22] In short, we should not expect a very high correlation between the amount of a grant and the percentage of cost or expenditure that it represents. In fact, the correlation between the amount of a grant and the percentage of tuition that it represents is only 0.251; the correlation between the amount of a grant and the percentage of comprehensive costs that it represents is 0.467; and the correlation between the amount of a grant and the percentage of per-pupil instructional spending that it represents is 0.310.

If students behave according a standard model of human capital investment, they should care about the *amount* of a grant, not the share of cost or expenditure that it represents. The amount of the grant is one of the two measures of the investment made freely *by others* in a student's human capital. (The other measure is the implicit grant given by the difference between a college's per-pupil expenditure and its "list" tuition. The size of a college's implicit grant increases monotonically with a college's selectivity. We have already seen that students appear to ignore per-pupil expenditure but respond strongly to a college's selectivity, suggesting that students probably do respond to implicit grants but do so using measures readily observable to them, such as selectivity.)

In any case, once a student knows the size of the grant that a college is offering him, it should be irrelevant to him what *share* of costs or expenditure it represents. Given the amount, he will care about the *share* only if he is an irrational investor (for instance, naively flattered by receiving a large *share* of tuition) or severely credit constrained (willing to give up others' large donations to his human capital investment in order to avoid having to make any cash contribution to that investment himself). Even if we suppose that colleges with high instructional spending are less efficient than others, it is unlikely their efficiency is so poor that 50 percent of a $14,000 expenditure ($7,000) truly represents a smaller investment than 100 percent of a $3,500 expenditure.

B. Evidence on how Students Respond to the Aid Variables They Should Ignore

Table 9 presents our estimates of how students make college choices when we allow they to respond to aid variables they "should ignore." Column I of the table merely reproduces the estimates in Table 3, in which students can respond to the amount of grants, but not their names, front-loading, or share of costs. The estimates suggest that each additional $1,000 of grant is estimated to increase the probability of matriculation by 11 percent of the prior probability.

---

[22] These include many flagship public universities of the South, Southwest, West (Wyoming, Utah, Colorado), and the California State University system (not the University of California). However, most of the low tuition group is made up of the least selective colleges to which students in the sample applied–branch campuses of public universities, non-flagship public colleges, and some low selectivity private colleges (which had a mean SAT percentile of 53, as compared to the average of 84 among colleges in the sample).

In column II, we add the indicators for the grant's being called a scholarship and being front-loaded. We also add variables indicating the share of tuition and the share of comprehensive cost that the grant represents. The first thing to note about the results in column II is that students no longer respond to the *amount* of the grant. The estimated effect of each additional $1,000 of grants is statistically insignificant, and the point estimate is not even greater than one. Loans, work-study, and other determinants of college choice have about the same effects that they had when we excluded the variables that "should" be ignored.

Column II also shows that, for a grant of a given amount, calling it a scholarship increases the probability of matriculation by 63.2 percent of the prior probability. This is a great effect for of an essentially hollow feature of a grant that any college could replicate at no cost. It is implausible that the indicator for a named scholarship is picking up part of the missing effect of grant amount: recall that the indicator is negatively correlated with the amount of the grant.

Front-loading also engenders a strong, positive matriculation response. For a grant of a given annual amount, its being front-loaded raises the probability of matriculation by 64.5 percent of the prior probability. Again, this is a substantial effect for a feature that costs a college little for students who are very likely to stay enrolled for four years (as are all of the students in the sample). Because of discounting and inflation, front-loading does cost a college something, but not much in comparison to the cost of inducing a student to attend by raising the amount of his grant. For instance, recall the annual grant of $4,000 and the front-loaded version mentioned above ($10,000 for one year, and $2,000 for three years). Discounting future years' spending at 6 percent, it costs the college an additional $654 to give the student the front-loaded version. However, in order to induce the same matriculation effect by raising the amount of the grant, the college would have had to raise the grant by $5,864 per year. While we should not take these numbers literally (since they are require a good deal of extrapolation and are based on different columns of Table 9), it is fairly obvious that students respond excessively to front-loading, as compared to the amount of the grant.

The next two rows of Column II show that students ignore the share of tuition that the grant represents, but place a great deal of weight on the share of comprehensive cost that the grant represents. (If we were to exclude the share of comprehensive cost, the share of tuition would–for obvious reasons–pick up much of the same effect. However, the share of comprehensive cost consistently explains much more of college choice than the share of tuition.) For every increase of 10 percent (0.10) in its share of comprehensive cost, a grant induces a student to raise his probability of matriculation by 207.3 percent of his prior probability. It is perfectly possible to have odds ratios of this size. For instance, if the student's prior probability of matriculation were 10 percent, his posterior probability would be 30.7 percent with a grant of 10 percent of comprehensive costs and would be 61.5 percent with a grant of 20 percent of comprehensive costs. What the high odds ratios are telling us, however, is that students are offered

grants that represent a large share of comprehensive costs only by colleges that they would have had low prior probabilities of attending in the absence of such grants. In addition, we should not really focus exclusively on the odds ratio on the share of comprehensive cost because the point estimate of the odds ratio on the share of tuition and the grant amount are below 1, and we know that the three variables are correlated.

Column III simply repeats the specification of column II, using only the sub-sample of observations for which we have data on per-pupil instructional spending. As expected, the statistical significance of most of the estimates fall, but most estimates are in the same range as in column II. However, the effect of a named scholarship falls, the effect of front-loading rises, and the effect of the grant's share of comprehensive cost rises substantially. The purpose of column III is merely to prepare us for column IV, which adds the college's per- pupil instructional spending and the grant's share of per-pupil instructional spending. Neither has a statistically significant effect on college choice.

Why is it that the grant's share of *comprehensive cost* matters, rather than its share of tuition or per-pupil instructional spending? Previously, we argued that a student would care about the grant's *share* only if he were an irrational investor (flattery) or severely credit constrained (unwilling or unable to pay cash). The grant's share of instructional spending is only weakly related to either motive. Instructional spending is poorly measured and difficult for students to observe, so the grant's share of instructional spending is unlikely to be sufficiently salient to be flattering. Also, the grant's share of instructional spending tells us little about a student's out-of-pocket payments because some colleges' spending substantially exceeds their "list" tuition and comprehensive cost. Students may respond to the grant's share of comprehensive cost rather than its share of tuition because comprehensive cost is more salient. After all, there is well-established term for getting a grant equal to 100 percent of comprehensive cost: "a free ride." There is no similarly accepted term for getting 100 percent of tuition. Also, the colleges that tend to offer grants that are large shares of comprehensive cost generally have low, even very low tuition ($1,500, $2000, $2,500, etc.), so that the non-tuition part of comprehensive cost is three to four times as large as tuition and typically makes up the lion's share of out-of-pocket expenses. Put another way, students show at least *some* sensitivity to the amount of the grant by responding more to the grant's share of comprehensive cost than to its share of tuition. Recall that, although neither correlation is very high, the grant's share of comprehensive cost is significantly more correlated with the amount of the grant than is the grant's share of tuition.

C. Do Students' Responses to Aid Variables that "Should Not Matter" Depend on their Backgrounds?

In Appendix Tables 3 through 5, we investigate whether a student's response to aid variables that "should not matter" depends on his background. We find some evidence that it does. To create the tables, we re-estimate the regression shown in Table 9 separately for students by parents' income group, parents' college selectivity group, and public versus private high school. Although we estimate odds ratios for all of the variables shown in Table 9,

Appendix Tables 3 through 5 present only the the selected coefficients relevant to our discussion of aid variables that "should not matter."

We find that a grant being called a scholarship significantly attracts students in every group, *except* students whose parents have high incomes or whose parents attended highly selective colleges. Also, we find that the grant's share of comprehensive cost attracts students who parents attended low and medium selectivity colleges, but *not* students whose parents attended high selectivity colleges. The latter students respond only to the amount of the grant.

Interestingly, when we investigate the very large average effect of the grant's share of comprehensive cost, we find that its significance depends crucially on students with medium-low parent incomes. Although the odds ratios for the other income groups are large also, the medium-low income group has by far the largest odds ratio and the only one that is statistically significant. This makes sense. Given their combination of merit and need, the low income students in our survey are eligible for and attract need-based aid. Their out-of-pocket contributions are small, even at expensive private colleges. At the other extreme, medium high and high income students apply less often to colleges that offer grants that are a large share of (a perhaps modest) comprehensive cost. In contrast, consider students whose parents have medium-low incomes. They are sufficiently well off to be asked for out-of-pocket payments that are substantial (though not nearly as those that would be asked of medium-high and high income students). Yet, the parents in the medium-low income group appear to vary greatly in their willingness or ability to pay out-of-pocket college expenses. In short, the circumstances of at least some medium-low income students generate maximum susceptibility to small grants that are represent a share of comprehensive costs.

We find that students from private and public high schools respond very similarly to aid variables that "should be ignored."

## IX. Are Students Making Reasonable Trade-Offs? Are They Too Attracted or Not Sufficiently Attracted by Aid?

We began with the project of determining whether students were using college aid as rational investors in the their own human capital. That is, are students making the best use of aid in order to maximize their lifetime economic well-being (which includes leisure, earnings from work, and other sources of income including aid)? Thus far, we have found two fairly obvious types of violations to rational human capital investment: students' responding excessively to loans and work-study, given their response to grants; and students' responding to aspects of grants that should not matter, for a grant of a given amount. However, both of these tests are essentially *relative*. That is, we have tested whether, *given his response to grants*, the student responds excessively to loans and work-

study. We have tested whether, *given his response to the grant amount*, the student responds excessively to aspects of the grant that should not affect his human capital investment decision.

We have not yet addressed our principal question: whether the student's response to aid is too great or too small in an *absolute* sense. Before addressing this question, we should note that it presupposes that some trade-off exists–that is, that in order to get more aid, a student must give something up. For instance, suppose a given student could be admitted to colleges A, B, and C, and that college C was preferable or at least as good as the others on the grounds of selectivity, the resources available for students, tuition (that is, lower tuition), campus life, location, and so on. Suppose also that college C offered more aid (that is, offered an aid package containing subsidies with a greater total value). Then, no trade-off would exist; the student would do better all around by matriculating at college C.

This no-trade off situation is unlikely to hold generally because it would not be an equilibrium. If all students saw the clear advantages of C, and received more aid, college C would be so over-subscribed that it would automatically become more selective, so that the student admitted to A and B might not longer be admitted to C. However, although it is useful to demonstrate logically why the situation would not be an equilibrium, it is also useful to show empirically that the situation does not exist, as a rule. This is straightforward. In the College Admissions Project survey data, we find that if we look *within* students' choice sets (so that we are holding student merit constant), the correlation between the grant amount that a student receives from a college and the college's median SAT score is -0.32. The corresponding correlation between the grant amount and the college's instructional expenditure is -0.36. These correlations suggest that, in general, a student must give up some college selectivity and/or some college resources in return for a larger grant. Put more bluntly, a student must generally allow the investment in his college education to be reduced in return for getting greater aid.

Of course, the fact that students are *generally* faced with a trade-off when they compare two colleges does not mean that students are *always* faced with one. For reasons that are idiosyncratic to the match between a particular student and college, a student may get the largest grant at the college at which he gets the most resources and which is the most selective. We have no difficulty with such idiosyncratic situations; we merely argue that such situations cannot be general as a logical matter and are not general as an empirical matter.

A. A Empirical Strategy for Determining whether Students Respond Too Much or Too Little to Aid

Returning to our principal question (whether the student's response to aid is too great or too small in an absolute sense), recall that the student is responding as a rational human capital investor if he makes the trade-off according to a condition such as the following:

(a) The subsidy value of the aid allows more consumption now, in return for an equally valuable decrease in future consumption, which will be caused by the reduced human capital investment.

(b) The subsidy value of the aid allows the students to make less use of loans, thereby reducing future interest payments in return for an equally valuable loss of future earnings, which will be reduced because of the decrease in human capital investment.[23]

Clearly, the key to addressing this question is measuring the extent to which student reduce their human capital investment in return for increased aid. In order to get such a measure, we use the estimated college choice model from the previous sections to perform some thought experiments. We take away all grants, loans, and other forms of aid (we zero out all of the aid variables), and we see how students' college choices are predicted to change. That is, we see what students would have done in the absence of aid. Our measures of their response to aid will be based on the difference between what they do with aid and what they would have done in the absence of it.

Once we have measures of how students's college choices change with aid, we will be able to estimate the losses associated with the reductions in human capital and consumption that they accept. We can then compare these losses to the value of the aid.

B. Measuring Human Capital Investment at a College

We would like to measure the human capital investment made in students at various colleges. In principle, there are two ways to do this, but we reject one as impracticable. First, we can measure human capital investment by attempting to measure the *inputs* available at a college (instructional resources and peers). Second, we can measure human capital investment by attempting to measure the *earnings outcomes* associated with a college and controlling for initial differences in students. We do not want to attribute all of a student's earnings to his college; much of his earnings are probably due to the abilities that got him admitted to the college in the first place.

Although either the measured-inputs or measured-outputs strategy could be pursued in theory, we, after serious consideration, reject the measured outputs strategy as impractical. There have been several empirical attempts to measure the value-added associated with various colleges, using students' later earnings. There is no set of commonly accepted estimates that we might use, and certain highly touted estimates are so flawed that we must conclude that readers are uninterested in the empirical details that make such a study credible or not. However, we do not reject the measured-outputs strategy merely to avoid controversy or having to present readers with empirical detail. We would be quite willing to generate a set of value-added estimates for this paper, were it possible. Unfortunately, one statement on which nearly every researcher could agree is that we are unable to estimate value-added for a *wide range* of *specific* colleges. One might argue that we can estimate value-added for a tiny number of specific colleges (the 32 colleges include in the *College and Beyond* survey); one might also argue that we can

[23] An additional possible trade-off, which we cannot observe in our data, works as follows. The grant might allow the student to work less and study more, thereby losing income in college but gaining income later in life because increased study is increased human capital investment.

estimate value-added for coarse groups of colleges (groups so coarse that some would include hundreds of colleges). No one would argue that we are able to estimate value-added for many, specific colleges. This is for a simple reason. There is no source of data that includes earnings and college identifiers for a broad array of colleges *and* has more than a few observations for any one college.[24] We need measures of human capital investments for nearly all of the colleges in our study if we are to determine whether students are making decisions like rational investors. We must use college-specific measures, not measures for coarse college groups: many of the students in our sample would be choosing within a single coarse group.

Therefore, we focus on the measured-inputs strategy for measuring human capital investment. Here, we are dealing with data that, though imperfect, are orders of magnitude more precise than that we could obtain using the measured-output strategy. Our approach to measuring inputs is very conservative and intentionally somewhat biased against high spending private colleges and high spending public research universities. We count instructional spending and *only* instructional spending as human capital investment. By focusing on instructional spending, we "bend over backward" in favor of finding human capital investment at low spending colleges that do not attempt to produce "complex" university services. Empirically, instructional spending is a *much* smaller share of total spending at high spending colleges than it is at low spending colleges. Among the colleges in our sample, there is a -0.32 correlation between instructional spending's share of total spending and total per-pupil spending. The colleges with highest total per-pupil spending have instructional spending shares around 0.1; the colleges with the lowest total per-pupil spending have instructional spending shares around 0.6. Moreover, instructional spending excludes spending on student services (such as health care), extracurricular activities designed to complement students' learning (such as a model United Nations), and research. These are important categories of expenditure at high spending colleges, which believe that they are complementary to instructional spending and thereby produce greater human capital. We treat all of these expenditures as though they were worthless.[25]

---

[24] The longitudinal surveys and the one Current Population Survey supplement that includes college identifiers have too few people in each college (often 0, 1, or 2); the *College and Beyond* survey includes numerous people in each college, but only includes a tiny group of colleges. The longitudinal surveys with college identifiers and a reasonably representative sample of the United States population are the *Panel Survey of Income Dynamics*, the *National Longitudinal Surveys*, and five surveys conducted by the United States Department of Education (the *National Longitudinal Study of the High School Class of 1972*, *High School and Beyond*, and the *National Education Longitudinal Study*, the *Beginning Postsecondary Student* survey and follow-up, and *Baccalaureate and Beyond*). The Current Population Survey supplement with college identifiers is the *Occupational Changes in a Generation* study (1972).

[25] We wish to clear up a common confusion, embodied in the following question: "Does not instructional spending understate the resources at a public college that is subsidized by the state?" Instructional spending does *not* understate resources at public colleges. State governments subsidize tuition and they often provide land and buildings below cost. The tuition subsidies are important for understanding the sources of *revenue* related to instruction, but they are irrelevant to instructional *spending*. We use instructional spending precisely because it is what it is, regardless of how the college's tuition is subsidized. Instructional spending excludes spending on

It is evident from the students' own choice behavior that they prefer more able peers, and it seems likely that part of the human capital gained by a student is generated by peer spillovers or by the interaction of good peer and college resources. We have no simple way to quantify peer-generated human capital at a college. Though we might attempt such quantification (and have, in fact, elsewhere), the attempt would require an elaborate empirical strategy beyond the scope of this paper. Nevertheless, we do not wish to ignore peer inputs and recognize only instructional inputs. Our "solution" to this problem is simply to show changes in peers' SAT scores and remind readers that they should keep peer quality in mind as an additional source of human capital, if they believe it to be important.

Although the students' choice behavior and common wisdom suggest that a student benefits from having more able peers, there is an alternative theory. A high aptitude student surrounded by significantly worse peers may be able to use much more than his share of a college's resources. While some attributes of a college must be shared relatively equally by all students (faculty quality), others (such as faculty time) can be disproportionately allocated to certain students. Logically, the extent of the disproportionateness must be a function of the degree to which a high aptitude students differs from his peers. If a high aptitude student attends a very selective college where he is typical, he cannot expect to receive much more than an equal share of the college's per-pupil resources.

In short, we show how a student's peers change when he accepts a certain college's aid package, but we leave readers to judge whether more able peers are net generators or destroyers of human capital investment, for a given level of college resources.

C. How Students' College Choices would Differ in the Absence of Aid, Part 1: The Structure of Table 10

To create Table 10, we estimate the conditional logit model shown in second column of Table 9, with one small exception. Because we eventually want to predict what college a student would choose, without aid, given a choice among all colleges (not just those in his actual choice set), it does not make sense to include the indicators for the college's being the most selective in his choice set and the least selective in his choice set. Instead, we rely more heavily on our two indicators of the match between the student's SAT score and the college's median SAT score. Recall that we included separate variables for positive and negative differences between the student's score and the college's median score. We now include not only these two separate variables, we also include square and the cube of each of the two variables. This specification helps us to account thoroughly for the change in the probability of matriculation as the match between the student's SAT score and the college's median score changes.

After estimating the conditional logit model, we use the coefficient estimates to predict which college the student would be most likely to attend in the absence of aid (that is, with all of the aid variables zeroed out). For simplicity, we will call the college that the student chose with actual aid his "with-aid college" and the college the student would have chosen without aid his "without-aid college." We then compute the value of the aid we zeroed

---

buildings and land, so all colleges are treated equally with regard to these two spending categories.

out, and we show this in column (1) of Table 10.[26] That is, column (1) shows the benefits of taking aid. The succeeding columns show the costs of taking aid. Column (2) shows the difference in consumption between the with-aid and without-aid colleges. We measure the difference in consumption by subtracting the room and board at the without-aid college from the room and board at his with-aid college. In other words, if the difference in consumption is negative, students at the without-aid college enjoy greater food and housing consumption than students at the with-aid college. (Recall, from the model of human capital investment, that we must account for the change in consumption. If all grants were merely tuition discounts, it might be appropriate to consider aid as one side of a student's trade-off and human capital investment (and only human capital investment) as the other side of the trade-off. But, we have seen that many grants are greater than tuition and only make sense in comparison to comprehensive costs. It would obviously be incorrect to count such tuition-exceeding grants on one side of the trade-off, yet exclude the consumption they finance from the other side of the trade-off.)

Column (3) shows the difference in instructional spending between the with-aid and the without-aid colleges. Column (4) shows the difference in the median SAT score, in percentiles, between with-aid and the without-aid colleges. This last measure is meant to indicate the change in a student's peer group.

There are two rows in Table 10. To generate these rows, we compute the lifetime benefits and costs of aid for each student, omitting the benefits and costs of peers, which we cannot measure accurately. We use a *real* discount rate of three percent per annum, assume an inflation rate of three percent, and assume a conservative seven percent real rate of return on human capital investment.[27] The lifetime value of a student's aid is simply the present value of four years of his aid package at his with-aid college. The lifetime value of the difference in a student's college consumption is the present value of four years of consumption at the with-aid college minus four years of consumption at the without-aid college. To get the lifetime value of the human capital acquired in college, we sum

---

[26] The vast majority of the variation in the value of aid comes from scholarships and other grants. Our results would not be noticeably affected by any reasonable procedure to estimate the subsidy value of loans and work-study commitments. In fact, we use federal estimates of the subsidy value of loans in the federal subsidized loan program. We do not attribute any subsidy value to loans with unsubsidized interest rates and repayment schedules. We assume that the subsidy value of a work-study commitment is one-third of its value.

[27] Only the present value of grants is affected by the inflation rate we choose because it is the only item that is generally fixed at the beginning of college. For instance, a National Merit Scholarship of $2,000 remains the same, in nominal dollars, through a student's college education. The expenditures on room, board, and instruction generally rise with inflation, as do loan limits and the work-study wage. A rate of return to human capital of 7 percent is near the bottom of the generally accepted range of estimates. It may be especially conservative for the highly meritorious group of students whom we are studying. Although there is no generally accepted set rate of return estimated especially for this group, more indicators suggest that they earn an unusually high rate of return on their baccalaureate years. See Card [2000] for a survey of rate of return to education calculations. Indeed, the tendency of highly meritorious students to continue in school beyond the baccalaureate degree strongly suggests that they earn a super-normal rate of return during their baccalaureate years, which prompts them to continue until their rate of return is more in line with their discount rate.

the instructional spending difference between the with-aid and without-aid colleges over four years of college. This gives us a human capital asset (which omits human capital associated with peers). We assume that the human capital asset begins paying out at seven percent after college (age 22) and continues paying out until age 65, at which time we assume retirement occurs.

Having made these calculations for each student, we divide students into two groups in order to summarize their response to aid. The two groups correspond to the two rows of Table 10, The first group (and row) contains students who, when we conducted our thought experiment, appeared to have made good use of the aid they were offered. Their original college choices *with aid* have higher lifetime values than they could achieve at their without-aid colleges *without-aid*. Keep in mind that, merely by sticking with the same college when aid is zeroed out, a student will be placed in this group; this is because his college variables will not change and the student's lifetime value will mechanically be higher because of the aid (it will be higher by an amount exactly equal to the present value of the aid itself). Because we could not account for the value of peers, this group's size is probably overstated but possibly understated. That is, there are some students who appear to have made good use of their aid but who might not have done so if peers generate human capital.

The second group (and row) contains students who, when we conducted our thought experiment, appeared to have responded excessively to the aid they were offered. That is, in return for aid, they accepted such large reductions in their human capital asset and consumption that they lost lifetime value. The size of this group is probably understated but possibly overstated because we cannot account for the value of peers. For instance, there may be students who appear to have given up more resources than they really have, because they enjoy a disproportionate share of their college's resources.

Of course, there is a third group of students whose without-aid college was the same as their with-aid college because they actually received no aid. It is not interesting to show changes for them because zeroing out their aid changes nothing: there is no "experiment" for them. Clearly, this group contains students who are not easily tempted by aid–if they were, they would presumably have attempted to get at least a few merit scholarships, which can obtained *at some college* by any student in our sample.

Table 10 has two panels. In the top panel, we show the differences between the with-aid and without-aid colleges when the student is allowed to re-choose among all of the colleges that appear in the College Admissions Project sample.[28] In the bottom panel, we show the differences between the with-aid and without-aid colleges when

[28] We think that it is reasonable to have the students choose just among the 755 colleges to which at least one surveyed student applied. This is because there are more than 3000 *other* institutions of higher education in the United States that virtually never enroll a student like those in our sample. They include community colleges, other two-year institutions, and other colleges that never or rarely grant the baccalaureate degree in a field of the humanities, social sciences, sciences, or engineering. Keeping these institutions as possible choices would simply slow our estimation, without adding explanatory power.

the student is restricting to re-choosing among the colleges in his initial choice set. The assumption that the student must re-choose within his initial choice set is inappropriately restrictive for our thought experiment, so we include the bottom panel only so that the reader may compute for himself how much application decisions are affected by aid. We will not discuss the bottom panel further.

D. How Students' College Choices would Differ in the Absence of Aid, Part 2: The Evidence in Table 10

The top panel of Table 10 shows that 34.9 percent of students responded to aid in such a way that their lifetime value was increased. At a minimum, they accepted aid at the same college they have chosen if no aid had been offered. More interestingly, many of the students in this group accepted an aid package that was sufficiently generous that it swamped the reduction in college consumption and human capital investment that they generally faced. Notice that the average value of aid for students in this group was high: $10,748 per year. They attended a with-aid college that offered consumption that was slightly lower ($283 per year), and that gave them $2,592 less per year in human capital investment. They probably also lost some human capital investment because their peers were slightly worse (3.5 percentile points worse on the SAT, relative to the peers they would have had at their without-aid college). On the other hand, their slightly worse peers may have allowed them to enjoy more than their share of resources. By responding like a rational investor to aid, the average student in this group gained lifetime present value of $48,923. This gain would be somewhat different if peers matter because since we have not deducted or added any amount for the student's worse peers.

Notice that the top row of Table 10 contains 34.9 percent of *all* students or 52.4 percent of students who received aid. It is important that the majority of students who actually received aid responded to it as we expect rational investors to respond. However, 31.7 of all students or 47.6 percent of students who received aid did *not* act like rational investors. The students in the bottow row of Table 10 accepted an aid package that was too small to make up for the losses in college consumption and human capital investment that they accepted. Notice that the average value of aid for students in this group was quite small: $2,754. Compared to the without-aid college they would have attended, students in this group attended a with-aid college that offered consumption that was $1,027 lower per year and that gave them $16,979 less per year in human capital investment. They probably also lost some human capital investment because their peers performed 8.8 percentile points worse than they did on the SAT. The average student in this group lost lifetime present value of $81,081. The loss would be somewhat different if peers matter, since we have not deducted or added any amount for the student's worse peers.

Readers may be initially surprised that a non-negligible minority of students lose so much when they respond to aid. But, the statistics in Table 10 are really a straightforward implication of the behavior that we saw illustrated in Table 9. The students who lose the most are precisely those students who accept aid that is actually quite modest in value because they respond to the fact that it covers a large share of comprehensive cost at a college

that spends very little on instruction. We know from Appendix Tables 3 and 4 that not all students are equally likely as losers of lifetime value: students who have high income parents or parents who are graduates of selective colleges themselves do not appear to be tempted by grants that are large shares of comprehensive costs at low spending colleges. These students react only to the actual amount of a grant. We suspect that these students behave more like rational investors either because they are more sophisticated and informed than other students or because they are less credit constrained than other students.

## X. Conclusions

Overall, we would describe the college choice behavior of the high aptitude students in our sample as systematic and sensitive to college attributes in the expected direction. We find that high aptitude students are nearly indifferent to a college's distance from their home or whether it is in-state or public. However, they care about less superficial college characteristics: they are sensitive to tuition, room, and board in the expected direction (lower is better), and they prefer to attend the most selective colleges in the set to which they are admitted. They are attracted by grants, loans, and work-study commitments. Although we find that students from different backgrounds do exhibit somewhat different college choice behavior, the differences are generally not dramatic and much college choice behavior is shared by the entire array of high aptitude students. The main exceptions to this rule are students whose parents have high incomes or who themselves graduated from very selective colleges. Such students exhibit less sensitivity to variables that affect college costs.

This being said, the students in our sample exhibit some hard-to-justify responses to grants and aid that they are offered. They are excessively attracted by loans and work-study, given the value of these types of aid compared to grants. They are attracted by superficial aspects of a grant, like its being called a scholarship (with a name) and its being front-loaded. They are far more sensitive to a grant's share of the college's comprehensive costs than they are to the amount of the grant. All these behaviors are deviations from the expected behavior of a rational investor in human capital. We should note that these peculiar behaviors are generally *not* shared by the students whose parents have high incomes or who themselves attended very selective colleges.

When we quantify the effect of students' responses to aid, we find that the majority of students (and a slim majority of actual aid recipients) respond as rational investors would in the presence of aid. They improve their lifetime present value by only accepting aid offers that are more than generous enough to offset the reductions in college consumption and human capital investment associated with the aid. However, a significant minority of students in our sample respond to aid in such a way that they reduce their own lifetime present value. Apart from idiosyncratic irrationality, there are two major, possible explanations for their behavior: a lack of sophistication and credit constraints.

A lack of sophistication accounts for at least some of the self-defeating responses to aid: credit constraints cannot explain why a student would be strongly attracted by a grant's being called a scholarship (when it costs a college nothing to do it). A lack of sophistication probably also accounts for the attraction of front-loaded grants–an alternative explanation is impatience, but this seems unlikely in a population of students who so obviously do not exhibit impatience (or any variant of hyperbolic discounting) as a rule. They all have records that show that they can work hard now in return for gains in the distant future. Credits constraints are also not a good explanation for the attractiveness of front-loading, since a front-loaded grant does not reduce the credit needs of families who know that their child will be enrolled for four years.

Either a lack of sophistication or credit constraints could explain the great attractiveness of grants that are a large share of comprehensive cost, regardless of that comprehensive cost is. It would probably be impossible to parse the effect into the share due to naivete and the share due to credit constraints. However, we did examine the open-ended comments by parents whose children exhibited the most self-defeating responses to aid. The overwhelming impression is that a lack of sophistication, and not credit constraints, are the problem. Over and over, these parents complain that they are baffled by the aid process. They plead for simpler forms and simpler offers. They argue that the colleges do not explain their alternatives well. They complain that the aid process requires financial records that they did not expect to need and cannot readily assemble. They complain that other families, especially those who use early applications, are more "in the know" and therefore have better outcomes. Only 6.9 percent of parents who commented and whose children exhibited self-defeating responses to aid stated that they were unable or unwilling to pay as much as their child's preferred college cost. We do not want to over-interpret the anecdotal evidence from parents' comments because they may be embarrassed to say that family circumstances prevent them from paying college costs. Nevertheless, we think that it is revealing that words like "bewildering" and "confusing" are the modal words in their comments.

**Econometric Appendix**

This appendix deals with three closely related issues in the econometrics of conditional logit estimation: the independence of irrelevant alternatives, endogenous choice sets, and unbalanced choice sets.

The conditional logit specification implies that the ratio of likelihoods of any two alternatives $j$ and $j'$

$$\frac{Prob(collegechoice_i = j)}{Prob(collegechoice_i = j')} \tag{5}$$

is independent of the other alternatives in the choice set. This property, the independence of irrelevant alternatives, is violated in certain applications. A stylized example may help. Suppose that people are choosing between two modes of getting to work: bus and bicycle. Suppose that the ratio of the bus likelihood to the bicycle likelihood is 1/3. Suppose that we add an alternative: a bicycle with racing stripes. Most people would probably a choose a bicycle with or without racing stripes without making reference to their bus-versus-bicycle decision. Yet, such behavior violates the independence of irrelevant alternatives. When the racing stripe bicycle is added to the choice set, some individuals' likelihood ratio for the bus versus the *plain* bicycle would shift because they would choose the racing stripe bicycle over the plain bicycle. Conditional logit estimation implies that the ratio of the bus likelihood to plain bicycle likelihood should be independent of whether a racing stripe bicycle is in the choice set. Conditional logit estimates are inconsistent in the presence of serious violations of the independence of irrelevant alternatives.

What do applied econometricians do about ensuring that their application satisfies the independence of irrelevant alternatives? There is a specification test for the determining whether a violation has taken place: Hausman and McFadden (1984), with alternative implementation suggested by McFadden (1987). It is a typical Hausman-type test in which, under the null hypothesis that likelihood ratios are indeed independent of irrelevant alternatives, excluding some alternatives from the choice set will produce inefficient but consistent estimates. Under the alternative hypothesis, consistency and not merely efficiency will be affected. Our application never comes close to being rejected by this specification test: for instance, the test statistic, which is distributed as $\chi^2_1$, is 0.43 for our basic specification in Table 3. It therefore has a p-value of .48. The reason our application never comes close to being rejected is fundamental and will become clear below.

Specification tests can only reveal that an assumption is violated. In order to ensure consistent estimates, econometricians often lump closely related alternatives together as a single alternative, creating the nested conditional logit model. For instance, in the bicycle example described above, both bicycles would be grouped together and the choice of a bicycle would be *prior* to the choice of bus versus bicycle. A difficulty with such nested models is that the econometrician must impose a nesting structure, often on somewhat *ad hoc* grounds. For the

bicycle versus bus example, it is obvious that the plain and racing stripe bicycle ought to be lumped together, but, in our application, it would be hard to choose whether to nest colleges based on their size, location, type (liberal arts, comprehensive, engineering, *et cetera*), or other criteria.

The most preferred method for satisfying the independence of irrelevant alternatives is finding a person's endogenous choice set. That is, we want individuals to perform the prior steps of the nested decision process themselves, revealing to us their final choice set. This method has been shown to produce consistent and precise estimates, even when the same observed choices would have violated the independence of irrelevant alternatives had the econometrician been ignorant of the endogenous choice set.[29] For instance, estimates of the effects of park amenities on park choices have been demonstrated to be more consistent and precise when individuals reveal the set of state and local parks that they considered relevant in making their choice. In most applications, econometricians do not observe endogenous choice sets, and they are therefore reduced to jointly estimating the endogenous choice set and the choice within the set. This is possible, as shown by Manski (1977), but it is very difficult to achieve identification in practice unless there is auxiliary evidence with which to predict a person's endogenous choice set. We are, however, in an almost uniquely fortunate situation. We observe true endogenous choice sets for nearly all of our student respondents. For the 10 percent of students who apply early decision and apply to only one college, we observe the choice sets of other similar students from the same high school. Thus, we either observe the endogenous choice set or can predict it with unusual accuracy. Fundamentally, this is why our data are never rejected by specification tests for the independence of irrelevant alternatives: by applying to college, students eliminate irrelevant alternatives for us and reveal their true endogenous choice sets.

With endogenous choice sets, different students have choice sets that contain different numbers of colleges and arrays of colleges (that is, the choice sets are "unbalanced"). Occasionally, we are asked whether this is a problem for consistency or whether it implies that students who have more colleges in their choice sets exercise disproportionate influence over the estimates. The answer is "no" on both counts, and questions themselves reveal a misunderstanding of conditional logit estimation. This becomes clear if we return to the conditional logit equations:

$$\ln L = \sum_{i=1}^{n} \left[ \sum_{j=1}^{J_i} matric_{ij} \ln Prob(collegechoice_i = j) \right] , \quad Prob(collegechoice_i = j) = \frac{e^{\beta' x_{ij}}}{\sum_{j=1}^{J_i} e^{\beta' x_{ij}}} .$$

Examine the log likelihood equation first. It shows that each student contributes equally to the log likelihood

---

[29] For proofs and examples of endogenous choice sets in conditional logit and multinomial logit models, see Manski (1977), Peters, Adamowicz, and Boxall (1995), Haab and Hicks (1997), Hicks and Strand (2000), and Parsons, Plantinga, and Boyle (2000). Manski's treatment is the most general, but he assumes that endogenous choice sets are not observed. He is therefore focused on how they may be estimated: identification is difficult, as it is in most models in which a prior selection stage must be estimated. The other papers listed focus on proofs and practice in circumstances where at least some auxiliary information is available on endogenous choice sets.

because each student's college choice probabilities must sum to one (see second equation). Therefore, the term inside the square brackets must be between 0 and 1 for each student. The equations also show why conditional logit is sometimes called fixed effects logit: the $\beta s$ only determine a college's *relative* probability within a student's choice set. Moreover, by choosing endogenous choice sets and eliminating irrelevant alternatives, the students help ensure that the assumptions imbedded in the second equation are true. That is, they help ensure that their choice behavior really does depend on the characteristics of the $J_i$ colleges that they consider. If there were irrelevant colleges included in their choice set, their behavior would contribute to estimation of $\beta s$ crucial to correct likelihood ratios of pairs of colleges *neither* of which is relevant to them.

**References**


Avery, Christopher, and Caroline M. Hoxby, "The College Admissions Project: Counselor Report," Harvard University typescript, 2000.

Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg, "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings," *Journal of Human Resources*, Vol. 34 (1), pp. 104-23. Winter 1999.

College Entrance Examination Board, *Standard Research Compilation: Undergraduate Institutions*. Electronic data, 2002.

Dynarski, Susan, "The Shift to Public Merit Aid and Its Effects," in Caroline M. Hoxby, ed., *College Decisions: How Students Actually Make Them and How They Could*, NBER typescript, 2002.

Ehrenberg, Ronald, "Optimal Financial Aid Policies for a Selective University," *Journal of Human Resources*, Vol. 19(2), 1984, pp. 202-230.

Haab, Timothy, and Robert Hicks, "Accounting for Choice Set Endogeneity in Random Utility Models of Recreation Demand," *Journal of Environmental Economics and Management*, Vol. 34(), 1997, pp. 127-147.

Hausman, Jerry, and Daniel McFadden, "A Specification Test for the Multinomial Logit Model," *Econometrica*, Vol. 52 (), 1984, pp. 1219-1240.

Hicks, Robert, and Ivar Strand, "The Extent of Information: Its Relevance for Random Utility Models," *Land Economics*, Vol. 76(3), 2000, pp. 374-385.

Hoxby, Caroline M., "The Return to Attending a More Selective College: 1960 to the Present," Harvard University typescript, 1998.

Hoxby, Caroline M., and Bridget Terry Long, "Explaining Rising Income and Wage Inequality among the College-Educated," NBER Working Paper No. 6873.

Kane, Thomas, "The Effects of Expectations About College: The Boston COACH Program and Beyond," in Caroline M. Hoxby, ed., *College Decisions: How Students Actually Make Them and How They Could*, NBER typescript, 2002.

Long, Bridget, "The Impact of the Federal Tax Credits for Higher Education Expenses," in Caroline M. Hoxby, ed., *College Decisions: How Students Actually Make Them and How They Could*, NBER typescript, 2002.

Manski, Charles, "The Structure of Random Utility Models," *Theory and Decision*, Vol. 8(), 1977, 229-254.

McFadden, Daniel, "Regression Based Specification Tests for the Multinomial Logit Model," *Journal of Econometrics*, Vol 34 () 1987, pp. 63-82.

Parsons, George, Andrew Plantinga, and Kevin Boyle, "Narrow Choice Sets in a Random Utility Model of Recreation Demand," *Land Economics*, Vol. 76(1), 2000, pp. 86-99.

Peters, Thomas., Wiktor Adamowicz, and Peter Boxall, "The Influence of Choice Set Consideration in Modeling the Benefits of Improved Water Quality," *Water Resources Research*, Vol. 613(), 1995, 1781-1787.

Peterson's, *Peterson's Guide to Four-Year Colleges*, 2002 edition. Princeton, NJ: Peterson's.

United States Department of Education, National Center for Education Statistics, *College Opportunities Online*. Electronic data, 2002.

United States Department of Education, National Center for Education Statistics, *Integrated Postsecondary Education Data System, Higher Education Finance Data File*. Electronic data, 2001.

Table 1
Description of the Students in the College Admission Project Data

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| Male | 0.4120 | 0.4923 | 0 | 1 | 3240 |
| White non-Hispanic | 0.7321 | 0.4429 | 0 | 1 | 3061 |
| Black | 0.0350 | 0.1837 | 0 | 1 | 3061 |
| Asian | 0.1571 | 0.3640 | 0 | 1 | 3061 |
| Hispanic | 0.0382 | 0.1918 | 0 | 1 | 3061 |
| Native American | 0.0010 | 0.0313 | 0 | 1 | 3061 |
| Other race/ethnicity | 0.0366 | 0.1878 | 0 | 1 | 3061 |
| Parents are married | 0.8305 | 0.3752 | 0 | 1 | 3228 |
| Sibling(s) enrolled in college | 0.2327 | 0.4226 | 0 | 1 | 3240 |
| Parents' income, estimated if necessary | 119929.0000 | 65518.2100 | 9186 | 240000 | 3218 |
| Parents' income < $20k | 0.0221 | 0.1469 | 0 | 1 | 3218 |
| Parents' income $20-30k | 0.0379 | 0.1910 | 0 | 1 | 3218 |
| Parents' income $30-40k | 0.0301 | 0.1710 | 0 | 1 | 3218 |
| Parents' income $40-50k | 0.0398 | 0.1955 | 0 | 1 | 3218 |
| Parents' income $50-60k | 0.0497 | 0.2174 | 0 | 1 | 3218 |
| Parents' income $60-70k | 0.0594 | 0.2363 | 0 | 1 | 3218 |
| Parents' income $70-80k | 0.0690 | 0.2535 | 0 | 1 | 3218 |
| Parents' income $80-90k | 0.0522 | 0.2225 | 0 | 1 | 3218 |
| Parents' income $90-100k | 0.0855 | 0.2796 | 0 | 1 | 3218 |
| Parents' income $100-120k | 0.1495 | 0.3566 | 0 | 1 | 3218 |
| Parents' income $120-140k | 0.0923 | 0.2895 | 0 | 1 | 3218 |
| Parents' income $140-160k | 0.0771 | 0.2667 | 0 | 1 | 3218 |
| Parents' income $160-200k | 0.0761 | 0.2653 | 0 | 1 | 3218 |
| Parents' income $200+k | 0.1594 | 0.3661 | 0 | 1 | 3218 |
| Expected family contribution, estimated if necessary | 27653.4700 | 16523.9200 | 0 | 120000 | 3218 |
| Applied for financial aid? | 0.5946 | 0.4910 | 0 | 1 | 3234 |
| Finances influenced college choice? | 0.4114 | 0.4922 | 0 | 1 | 3053 |
| Amount of outside scholarships, applicable at *any* college | 203.0781 | 799.9640 | 0 | 12500 | 3240 |
| National Merit Scholarship winner | 0.0494 | 0.2167 | 0 | 1 | 3240 |
| Student's SAT score, sum of math and verbal, converted from ACT score if necessary | 1356.9110 | 138.8193 | 780 | 1600 | 3192 |
| Student's SAT score, expressed as national percentile | 90.4013 | 12.3362 | 12 | 100 | 3192 |
| Median SAT score at *most* selective college to which student was admitted | 86.4092 | 10.3836 | 34 | 98 | 3203 |
| Median SAT score at *least* selective college to which student was admitted | 73.8469 | 14.5646 | 14 | 97 | 3203 |
| Number of colleges to which student was admitted | 3.5250 | 2.1293 | 1 | 10 | 3240 |
| Student's high school was private | 0.4534 | 0.4979 | 0 | 1 | 3240 |
| Student's high school in AL | 0.0170 | 0.1292 | 0 | 1 | 3240 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| Student's high school in AR | 0.0028 | 0.0526 | 0 | 1 | 3240 |
| Student's high school in AZ | 0.0093 | 0.0958 | 0 | 1 | 3240 |
| Student's high school in CA | 0.1222 | 0.3276 | 0 | 1 | 3240 |
| Student's high school in CO | 0.0120 | 0.1091 | 0 | 1 | 3240 |
| Student's high school in CT | 0.0327 | 0.1779 | 0 | 1 | 3240 |
| Student's high school in DC | 0.0096 | 0.0974 | 0 | 1 | 3240 |
| Student's high school in FL | 0.0287 | 0.1670 | 0 | 1 | 3240 |
| Student's high school in GA | 0.0111 | 0.1048 | 0 | 1 | 3240 |
| Student's high school in HI | 0.0201 | 0.1402 | 0 | 1 | 3240 |
| Student's high school in ID | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in IL | 0.0633 | 0.2435 | 0 | 1 | 3240 |
| Student's high school in IN | 0.0086 | 0.0926 | 0 | 1 | 3240 |
| Student's high school in KS | 0.0046 | 0.0679 | 0 | 1 | 3240 |
| Student's high school in KY | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in LA | 0.0105 | 0.1019 | 0 | 1 | 3240 |
| Student's high school in MA | 0.0855 | 0.2797 | 0 | 1 | 3240 |
| Student's high school in MD | 0.0327 | 0.1779 | 0 | 1 | 3240 |
| Student's high school in ME | 0.0052 | 0.0723 | 0 | 1 | 3240 |
| Student's high school in MI | 0.0198 | 0.1392 | 0 | 1 | 3240 |
| Student's high school in MN | 0.0056 | 0.0743 | 0 | 1 | 3240 |
| Student's high school in MO | 0.0198 | 0.1392 | 0 | 1 | 3240 |
| Student's high school in MT | 0.0019 | 0.0430 | 0 | 1 | 3240 |
| Student's high school in NC | 0.0219 | 0.1464 | 0 | 1 | 3240 |
| Student's high school in NE | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in NH | 0.0167 | 0.1280 | 0 | 1 | 3240 |
| Student's high school in NJ | 0.0522 | 0.2224 | 0 | 1 | 3240 |
| Student's high school in NM | 0.0102 | 0.1004 | 0 | 1 | 3240 |
| Student's high school in NV | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in NY | 0.1278 | 0.3339 | 0 | 1 | 3240 |
| Student's high school in OH | 0.0309 | 0.1730 | 0 | 1 | 3240 |
| Student's high school in OK | 0.0062 | 0.0783 | 0 | 1 | 3240 |
| Student's high school in OR | 0.0105 | 0.1019 | 0 | 1 | 3240 |
| Student's high school in PA | 0.0472 | 0.2121 | 0 | 1 | 3240 |
| Student's high school in RI | 0.0086 | 0.0926 | 0 | 1 | 3240 |
| Student's high school in SC | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in TN | 0.0201 | 0.1402 | 0 | 1 | 3240 |
| Student's high school in TX | 0.0395 | 0.1948 | 0 | 1 | 3240 |
| Student's high school in UT | 0.0071 | 0.0840 | 0 | 1 | 3240 |
| Student's high school in VA | 0.0333 | 0.1795 | 0 | 1 | 3240 |
| Student's high school in VT | 0.0031 | 0.0555 | 0 | 1 | 3240 |
| Student's high school in WA | 0.0160 | 0.1257 | 0 | 1 | 3240 |
| Student's high school in WI | 0.0077 | 0.0875 | 0 | 1 | 3240 |
| Student's high school in WY | 0.0028 | 0.0526 | 0 | 1 | 3240 |

Table 2
Description of the Colleges to Which Students *Were Admitted*,
from the College Admission Project Data

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| Matriculated at this college | 0.2825 | 0.4502 | 0 | 1 | 11455 |
| Admitted to this college | 1.0000 | 0.0000 | 1 | 1 | 11468 |
| Applied early to this college | 0.1298 | 0.3405 | 0 | 2 | 11453 |
| Withdrew application from this college, usually after early decision elsewhere | 0.0000 | 0.0000 | 0 | 0 | 11453 |
| Grants specific to this college | 2719.8600 | 5870.0240 | 0 | 36000 | 11468 |
| Loans from this college | 641.3459 | 2282.1720 | 0 | 36548 | 11468 |
| Work study amount from this college | 172.1048 | 593.0736 | 0 | 15000 | 11468 |
| Grant is called a named "scholarship" | 0.1958 | 0.3968 | 0 | 1 | 11468 |
| Grant is front-loaded (more in freshman year) | 0.0212 | 0.1440 | 0 | 1 | 11468 |
| Grant is this share of tuition | 0.1885 | 0.4369 | 0 | 7 | 11388 |
| Grant is this share of comprehensive cost | 0.1109 | 0.2258 | 0 | 2 | 11380 |
| Student was a recruited athlete at this college | 0.0275 | 0.1634 | 0 | 1 | 11468 |
| Father is an alumnus of this college | 0.0401 | 0.1962 | 0 | 1 | 11468 |
| Mother is an alumna of this college | 0.0283 | 0.1659 | 0 | 1 | 11468 |
| Sibling attended or attends this college | 0.0484 | 0.2146 | 0 | 1 | 11468 |
| College is public | 0.3325 | 0.4711 | 0 | 1 | 11468 |
| College is private not-for-profit | 0.6628 | 0.4737 | 0 | 1 | 11468 |
| College is international, except for Canadian colleges which are treated as U.S. colleges | 0.0045 | 0.0672 | 0 | 1 | 11468 |
| College's median SAT score, in national percentiles | 80.5947 | 12.5188 | 14 | 98 | 10778 |
| Student's SAT score is this many percentiles *above* college's median SAT score | 11.2945 | 10.2160 | 0 | 82 | 10654 |
| Student's SAT score is this many percentiles *below* college's median SAT score | 1.1006 | 4.3038 | 0 | 58 | 10654 |
| In-state tuition | 16435.1500 | 9594.0020 | 0 | 27472 | 11429 |
| Out-of-state tuition | 19293.5700 | 6190.8330 | 0 | 27472 | 11429 |
| Tuition that applies to this student | 17670.6000 | 8491.8630 | 0 | 27472 | 11429 |
| Room and board at this college | 6808.9370 | 1322.2720 | 0 | 10299 | 11407 |
| In-state comprehensive cost of this college | 23785.2000 | 10368.3300 | 0 | 35125 | 11407 |
| Out-of-state comprehensive cost of this college | 26641.5400 | 7032.6210 | 0 | 35125 | 11407 |
| Comprehensive cost that applies to this student | 25022.2000 | 9219.1590 | 0 | 35125 | 11407 |
| Total per-pupil expenditure of this college, in thousands | 58.3288 | 54.6237 | 5 | 816 | 8222 |
| Educational and general per-pupil expenditure of this college, in thousands | 41.9290 | 24.4120 | 5 | 168 | 8222 |
| Instructional per-pupil expenditure of this college, in thousands | 15.2391 | 10.5614 | 2 | 51 | 8222 |
| College is in-state | 0.3270 | 0.4691 | 0 | 1 | 11468 |
| Distance between student's high school and this college, in miles | 597.1856 | 808.9188 | 0 | 5769 | 10958 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in AK | 0.0000 | 0.0000 | 0 | 0 | 11389 |
| College is in AL | 0.0053 | 0.0724 | 0 | 1 | 11389 |
| College is in AR | 0.0004 | 0.0187 | 0 | 1 | 11389 |
| College is in AZ | 0.0056 | 0.0748 | 0 | 1 | 11389 |
| College is in CA | 0.1385 | 0.3454 | 0 | 1 | 11389 |
| College is in CO | 0.0109 | 0.1038 | 0 | 1 | 11389 |
| College is in CT | 0.0380 | 0.1913 | 0 | 1 | 11389 |
| College is in DC | 0.0260 | 0.1591 | 0 | 1 | 11389 |
| College is in DE | 0.0032 | 0.0561 | 0 | 1 | 11389 |
| College is in FL | 0.0164 | 0.1271 | 0 | 1 | 11389 |
| College is in GA | 0.0197 | 0.1389 | 0 | 1 | 11389 |
| College is in HI | 0.0035 | 0.0592 | 0 | 1 | 11389 |
| College is in IA | 0.0042 | 0.0648 | 0 | 1 | 11389 |
| College is in ID | 0.0013 | 0.0363 | 0 | 1 | 11389 |
| College is in IL | 0.0543 | 0.2265 | 0 | 1 | 11389 |
| College is in IN | 0.0206 | 0.1422 | 0 | 1 | 11389 |
| College is in KS | 0.0022 | 0.0468 | 0 | 1 | 11389 |
| College is in KY | 0.0006 | 0.0248 | 0 | 1 | 11389 |
| College is in LA | 0.0094 | 0.0965 | 0 | 1 | 11389 |
| College is in MA | 0.1054 | 0.3070 | 0 | 1 | 11389 |
| College is in MD | 0.0219 | 0.1462 | 0 | 1 | 11389 |
| College is in ME | 0.0144 | 0.1191 | 0 | 1 | 11389 |
| College is in MI | 0.0227 | 0.1488 | 0 | 1 | 11389 |
| College is in MN | 0.0089 | 0.0938 | 0 | 1 | 11389 |
| College is in MO | 0.0259 | 0.1589 | 0 | 1 | 11389 |
| College is in MS | 0.0009 | 0.0296 | 0 | 1 | 11389 |
| College is in MT | 0.0010 | 0.0311 | 0 | 1 | 11389 |
| College is in NC | 0.0356 | 0.1852 | 0 | 1 | 11389 |
| College is in NE | 0.0018 | 0.0419 | 0 | 1 | 11389 |
| College is in NH | 0.0118 | 0.1078 | 0 | 1 | 11389 |
| College is in NJ | 0.0217 | 0.1457 | 0 | 1 | 11389 |
| College is in NM | 0.0017 | 0.0408 | 0 | 1 | 11389 |
| College is in NV | 0.0008 | 0.0281 | 0 | 1 | 11389 |
| College is in NY | 0.1212 | 0.3263 | 0 | 1 | 11389 |
| College is in OH | 0.0273 | 0.1630 | 0 | 1 | 11389 |
| College is in OK | 0.0018 | 0.0419 | 0 | 1 | 11389 |
| College is in OR | 0.0087 | 0.0928 | 0 | 1 | 11389 |
| College is in PA | 0.0713 | 0.2573 | 0 | 1 | 11389 |
| College is in RI | 0.0193 | 0.1376 | 0 | 1 | 11389 |
| College is in SC | 0.0049 | 0.0700 | 0 | 1 | 11389 |
| College is in TN | 0.0139 | 0.1170 | 0 | 1 | 11389 |
| College is in TX | 0.0222 | 0.1474 | 0 | 1 | 11389 |
| College is in UT | 0.0045 | 0.0668 | 0 | 1 | 11389 |
| College is in VA | 0.0391 | 0.1938 | 0 | 1 | 11389 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in VT | 0.0104 | 0.1013 | 0 | 1 | 11389 |
| College is in WA | 0.0122 | 0.1098 | 0 | 1 | 11389 |
| College is in WI | 0.0090 | 0.0942 | 0 | 1 | 11389 |
| College is in WV | 0.0000 | 0.0000 | 0 | 0 | 11389 |
| College is in WY | 0.0003 | 0.0162 | 0 | 1 | 11389 |

Table 3
The Determinants of College Choice
Estimated Odds Ratios from Conditional Logit Regressions in which the Binary Outcome is Matriculation

| | | |
|---|---|---|
| Grant (in thousands), specific to the college | **1.110**<br>**(15.890)** | **1.113**<br>**(13.170)** |
| Loan (in thousands) from the college | **1.098**<br>**(5.770)** | **1.108**<br>**(3.900)** |
| Work study amount (in thousands) from the college | **1.414**<br>**(5.460)** | **1.428**<br>**(4.150)** |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | **0.980**<br>**(-3.230)** | 0.975<br>(-1.730) |
| College's room and board (in thousands) | **0.909**<br>**(-3.190)** | 0.948<br>(-1.380) |
| College's per-pupil instructional spending (in thousands) | | 0.999<br>(-0.340) |
| Student's SAT score is this number of percentiles *above* college's average SAT score | **0.951**<br>**(-7.940)** | **0.954**<br>**(-5.480)** |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 1.003<br>(0.280) | 0.989<br>(-0.770) |
| College is *most* selective to which student was admitted | **1.762**<br>**(8.730)** | **2.112**<br>**(9.520)** |
| College is *least* selective to which student was admitted | **0.701**<br>**(-4.150)** | **0.778**<br>**(-2.260)** |
| Father is alumnus of this college | **1.715**<br>**(3.680)** | **1.474**<br>**(1.980)** |
| Mother is alumna of this college | 0.970<br>(-0.160) | 1.067<br>(0.230) |
| Sibling attended or attends this college | **1.949**<br>**(5.280)** | **1.612**<br>**(2.610)** |
| Distance between college and student's high school, in hundreds of miles | 1.000<br>(-0.010) | 0.984<br>(-1.850) |
| Square of distance between college and student's high school, in 10000s of miles | 1.000<br>(1.060) | 1.000<br>(1.350) |
| Cube of distance between college and student's high school, in 1000000s of miles | 1.000<br>(-1.050) | 1.000<br>(-1.300) |
| College is in-state for the student | 1.139<br>(1.410) | 0.998<br>(-0.020) |
| College is public | 1.096<br>(0.640) | 1.329<br>(1.210) |
| Number of observations | 9113 | 5513 |
| Likelihood ratio ($chi^2$) | 1098.36 | 709.31 |
| Prob>$chi^2$ | 0 | 0 |
| LogLikelihood | -2372.283 | -1465.84 |
| Pseudo $R^2$ | 0.188 | 0.195 |

Notes: The table shows results from conditional logit estimation of how a student chooses his matriculation college among the colleges to which he was admitted. The results are shown as odds ratios, with Z-statistics in parentheses below the odd ratios. Results shown in bold print are odds ratios that are statistically significantly different from 1 with at least 95 percent confidence. The source of data is the College Admissions Project.

Table 4
Including Early Decision Students in Estimates of College Choice,
Estimated Odds Ratios from Conditional Logit Regressions in which the Binary Outcome is Matriculation

| | actual choice sets | predicted choice sets used for early applicants |
|---|---|---|
| Grant (in thousands), specific to the college | **1.110**<br>**(15.890)** | **1.107**<br>**(15.650)** |
| Loan (in thousands) from the college | **1.098**<br>**(5.770)** | **1.101**<br>**(5.900)** |
| Work study amount (in thousands) from the college | **1.414**<br>**(5.460)** | **1.399**<br>**(5.270)** |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | **0.980**<br>**(-3.230)** | 0.985<br>(-1.590) |
| College's room and board (in thousands) | **0.909**<br>**(-3.190)** | **0.914**<br>**(-2.970)** |
| Student's SAT score is this number of percentiles *above* college's average SAT score | **0.951**<br>**(-7.940)** | **0.911**<br>**(-17.100)** |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 1.003<br>(0.280) | 0.985<br>(-1.240) |
| College is *most* selective to which student was admitted | **1.762**<br>**(8.730)** | **3.115**<br>**(9.260)** |
| College is *least* selective to which student was admitted | **0.701**<br>**(-4.150)** | **2.211**<br>**(4.850)** |
| Father is alumnus of this college | **1.715**<br>**(3.680)** | **1.688**<br>**(3.610)** |
| Mother is alumna of this college | 0.970<br>(-0.160) | 0.956<br>(-0.240) |
| Sibling attended or attends this college | **1.949**<br>**(5.280)** | **1.921**<br>**(5.240)** |
| Distance between college and student's high school, in hundreds of miles | 1.000<br>(-0.010) | 0.993<br>(-1.040) |
| Square of distance between college and student's high school, in 10000s of miles | 1.000<br>(1.060) | 1.000<br>(0.830) |
| Cube of distance between college and student's high school, in 1000000s of miles | 1.000<br>(-1.050) | 1.000<br>(-0.820) |
| College is in-state for the student | 1.139<br>(1.410) | 1.055<br>(0.580) |
| College is public | 1.096<br>(0.640) | 1.130<br>(0.860) |
| Number of observations | 9113 | 9194 |
| Likelihood ratio ($chi^2$) | 1098.36 | 1098.75 |
| Prob>$chi^2$ | 0 | 0 |
| LogLikelihood | -2372.283 | -2393.692 |
| Pseudo $R^2$ | 0.188 | 0.187 |

Notes: This table is the same as Table 3, except that the right-hand column substitutes predicted choice sets for actual choice sets for those students who apply early to a college, are accepted by that college, and do not apply elsewhere. The prediction procedure is described in Sections IV of the text. All other notes from Table 3 apply.

Table 5
Are Students from Low and High Income Families Equally Sensitive to the Determinants of College Choice?
Estimated Odds Ratios from Conditional Logit Regressions
(notes continue on next page)

| | Parents Income is: | | | | |
|---|---|---|---|---|---|
| | Low | Medium Low | Medium High | High | same effect for all groups? |
| Grant (in thousands), specific to the college | **1.126** **(6.230)** | **1.130** **(8.670)** | **1.132** **(11.060)** | **1.075** **(5.240)** | rejected |
| Loan (in thousands) from the college | **1.111** **(2.020)** | **1.121** **(3.080)** | **1.105** **(3.710)** | **1.086** **(2.940)** | |
| Work study amount (in thousands) from the college | 1.068 (0.390) | **1.567** **(3.980)** | **1.371** **(2.970)** | **1.692** **(2.710)** | rejected |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | 0.997 (-0.150) | **0.964** **(-2.510)** | **0.968** **(-3.180)** | 0.992 (-0.670) | rejected |
| College's room and board (in thousands) | 0.840 (-1.590) | 0.933 (-1.060) | 0.916 (-1.760) | 0.919 (-1.520) | |
| Student's SAT score is this number of percentiles *above* college's average SAT score | 0.971 (-1.720) | **0.940** **(-4.420)** | **0.948** **(-5.300)** | **0.948** **(-4.160)** | rejected |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 0.992 (-0.320) | 1.023 (0.950) | 1.010 (0.550) | 1.020 (0.720) | |
| College is *most* selective to which student was admitted | 1.525 (1.870) | **1.625** **(3.410)** | **1.573** **(4.210)** | **2.093** **(6.330)** | rejected |
| College is *least* selective to which student was admitted | 1.034 (0.120) | 0.785 (-1.270) | **0.638** **(-3.280)** | **0.681** **(-2.340)** | rejected |
| Father is alumnus of this college | 3.288 (1.530) | 1.360 (0.820) | **2.341** **(3.510)** | 1.301 (1.100) | rejected |
| Mother is alumna of this college | 0.221 (-1.140) | 2.006 (1.560) | 0.531 (-1.960) | 1.727 (1.790) | rejected |
| Sibling attended or attends this college | **2.592** **(2.140)** | **1.934** **(2.260)** | **2.309** **(4.250)** | 1.395 (1.390) | rejected |
| Distance between college and student's high school, in hundreds of miles | 0.973 (-1.100) | 1.004 (0.250) | 1.002 (0.170) | 0.997 (-0.210) | |
| Square of distance between college and student's high school, in 10000s of miles | **1.000** **(-2.420)** | 1.000 (0.530) | 1.000 (-0.090) | 1.000 (1.170) | rejected |
| Cube of distance between college and student's high school, in 1000000s of miles | **1.004** **(2.410)** | 1.000 (-0.530) | 1.000 (0.100) | 1.000 (-1.130) | rejected |
| College is in-state for the student | 0.927 (-0.240) | 0.973 (-0.130) | **1.386** **(2.170)** | 0.961 (-0.230) | |
| College is public | 1.621 (1.03) | 1.739 (1.760) | 1.031 (0.130) | 0.954 (-0.270) | |
| Number of observations | 838 | 2012 | 3459 | 2731 | |
| Likelihood ratio ($chi^2$) | 135.99 | 329.91 | 429.46 | 295.62 | |
| Prob>$chi^2$ | 0 | 0 | 0 | 0 | |
| LogLikelihood | -202.658 | -478.443 | -898.309 | -724.28 | |
| Pseudo $R^2$ | 0.251 | 0.256 | 0.193 | 0.17 | |

Notes: The table shows results from conditional logit estimation of how a student chooses his matriculation college among the colleges to which he was admitted. The results are shown as odds ratios, with Z-statistics in parentheses below the odd ratios. Results shown in bold print are odds ratios that are statistically significantly different from 1 with at least 95 percent confidence. Parents are divided into four income groups: low, less than $40,000; medium low, $40,000 to $80,000; medium high, $80,000 to $140,000; high, greater than or equal to $140000. The right hand column contains the word "rejected" when the hypothesis that the odds ratios for the four income groups are equal is rejected with 95 percent confidence. The source of data is the College Admissions Project.

Table 6
Are Students whose Parents Attended More and Less Selective Colleges Equally Sensitive to the Determinants of College Choice?
Estimated Odds Ratios from Conditional Logit Regressions
(notes continue on next page)

| | Parents' college selectivity is: | | | |
|---|---|---|---|---|
| | low | medium | high | same effect for all groups? |
| Grant (in thousands), specific to the college | **1.123**<br>**(14.240)** | **1.097**<br>**(6.490)** | **1.080**<br>**(3.500)** | rejected |
| Loan (in thousands) from the college | **1.111**<br>**(5.460)** | **1.121**<br>**(2.410)** | 1.041<br>(1.050) | |
| Work study amount (in thousands) from the college | **1.258**<br>**(3.000)** | **1.560**<br>**(3.380)** | **3.181**<br>**(3.700)** | rejected |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | **0.978**<br>**(-2.800)** | **0.953**<br>**(-3.590)** | **1.042**<br>**(1.960)** | |
| College's room and board (in thousands) | **0.886**<br>**(-3.150)** | 1.015<br>(0.250) | 0.851<br>(-1.730) | |
| Student's SAT score is this number of percentiles *above* college's average SAT score | **0.954**<br>**(-6.200)** | **0.935**<br>**(-4.580)** | **0.948**<br>**(-2.260)** | |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 0.999<br>(-0.060) | 1.029<br>(0.950) | **1.173**<br>**(1.940)** | rejected |
| College is *most* selective to which student was admitted | **1.611**<br>**(5.710)** | **2.029**<br>**(5.310)** | **1.706**<br>**(2.870)** | |
| College is *least* selective to which student was admitted | **0.699**<br>**(-3.340)** | **0.682**<br>**(-2.100)** | 0.705<br>(-1.350) | |
| Father is alumnus of this college | 1.210<br>(0.600) | **1.813**<br>**(2.560)** | 1.550<br>(1.600) | |
| Mother is alumna of this college | 1.360<br>(0.880) | 0.736<br>(-1.020) | 0.861<br>(-0.370) | |
| Sibling attended or attends this college | **1.966**<br>**(4.350)** | **2.632**<br>**(3.730)** | 0.844<br>(-0.380) | rejected |
| Distance between college and student's high school, in hundreds of miles | 1.010<br>(1.120) | **0.881**<br>**(-4.500)** | 1.038<br>(1.620) | rejected |
| Square of distance between college and student's high school, in 10000s of miles | 1.000<br>(0.630) | **1.000**<br>**(2.060)** | 1.000<br>(-0.010) | |
| Cube of distance between college and student's high school, in 1000000s of miles | 1.000<br>(-0.630) | 1.000<br>(-1.340) | 1.000<br>(-0.040) | |
| College is in-state for the student | 1.143<br>(1.160) | **0.680**<br>**(-1.960)** | **2.165**<br>**(2.690)** | rejected |
| College is public | 1.079<br>(0.430) | 1.500<br>(1.340) | 0.759<br>(-0.560) | |
| Number of observations | 5674 | 2280 | 1159 | |
| Likelihood ratio ($chi^2$) | 678.12 | 346.17 | 168.82 | |
| Prob>$chi^2$ | 0 | 0 | 0 | |
| LogLikelihood | -1475.689 | -563.883 | -285.335 | |
| Pseudo $R^2$ | 0.187 | 0.235 | 0.228 | |

Notes: The table shows results from conditional logit estimation of how a student chooses his matriculation college among the colleges to which he was admitted. The results are shown as odds ratios, with Z-statistics in parentheses below the odd ratios. Results shown in bold print are odds ratios that are statistically significantly different from 1 with at least 95 percent confidence. Parents are divided into college selectivity groups, based on the *maximum* selectivity of the two parents' colleges: low, college's median SAT is less than the 70th percentile; medium,college's median SAT is between the 70th and 90th percentile; high, college's median SAT is greater than or equal to the 90th percentile. Note that selectivity is based on colleges' current selectivity, owing to the paucity of data on selectivity for the years in parents attended college. The right hand column contains the word "rejected" when the hypothesis that the odds ratios for the three selectivity groups are equal is rejected with 95 percent confidence. The source of data is the College Admissions Project.

Table 7
Are Students from Private and Public High Schools Equally Sensitive to the Determinants of College Choice?
Estimated Odds Ratios from Conditional Logit Regressions
(notes continue on next page)

| | Student's high school is private or public? | | |
|---|---|---|---|
| | public | private | same effect for all groups? |
| Grant (in thousands), specific to the college | **1.113** **(12.040)** | **1.109** **(10.370)** | |
| Loan (in thousands) from the college | **1.112** **(5.200)** | **1.085** **(3.030)** | |
| Work study amount (in thousands) from the college | **1.485** **(4.840)** | **1.339** **(2.860)** | |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | 0.988 (-1.460) | **0.973** **(-2.650)** | |
| College's room and board (in thousands) | **0.824** **(-4.530)** | 1.001 (0.010) | rejected |
| Student's SAT score is this number of percentiles *above* college's average SAT score | **0.956** **(-5.650)** | **0.944** **(-5.540)** | |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 0.993 (-0.450) | 1.019 (1.060) | |
| College is *most* selective to which student was admitted | **1.829** **(6.700)** | **1.681** **(5.460)** | |
| College is *least* selective to which student was admitted | **0.796** **(-2.010)** | **0.558** **(-4.310)** | rejected |
| Father is alumnus of this college | **1.970** **(3.440)** | 1.500 (1.790) | |
| Mother is alumna of this college | 0.771 (-1.030) | 1.339 (0.970) | |
| Sibling attended or attends this college | **2.190** **(5.030)** | 1.517 (1.890) | |
| Distance between college and student's high school, in hundreds of miles | 1.002 (0.240) | 0.987 (-1.040) | |
| Square of distance between college and student's high school, in 10000s of miles | 1.000 (0.160) | 1.000 (1.180) | |
| Cube of distance between college and student's high school, in 1000000s of miles | 1.000 (-0.160) | 1.000 (-1.080) | |
| College is in-state for the student | **1.298** **(2.090)** | 0.955 (-0.330) | rejected |
| College is public | 0.880 (-0.690) | **1.792** **(2.510)** | rejected |
| Number of observations | 4818 | 4295 | |
| Likelihood ratio ($chi^2$) | 620.49 | 516.81 | |
| Prob>$chi^2$ | 0 | 0 | |
| LogLikelihood | -1253.654 | -1099.157 | |
| Pseudo $R^2$ | 0.198 | 0.19 | |

Notes: The table shows results from conditional logit estimation of how a student chooses his matriculation college among the colleges to which he was admitted. The results are shown as odds ratios, with Z-statistics in parentheses below the odd ratios. Results shown in bold print are odds ratios that are statistically significantly different from 1 with at least 95 percent confidence. Students are divided into two groups, based on the control (private or public) of their high schools. The right hand column contains the word "rejected" when the hypothesis that the odds ratio for the two groups are equal is rejected with 95 percent confidence. The source of data is the College Admissions Project.

Table 8
The Share of Students who would Matriculate at Another College
If They were to Obey the College Choice Model of Students from a Different Background
Predictions Based on Conditional Logit Regressions Shown in Tables 4 through 6

| | Students whose Parent Income is: | | Students whose Parents' College Selectivity was: | | Students whose High School was: | |
|---|---|---|---|---|---|---|
| | Actually Low, Acting Like High Income Students | Actually High, Acting like Low Income Students | Actually Low, Acting as though it had been High | Actually High, Acting as though it had been Low | Actually Public, Acting like Private School Students | Actually Private, Acting like Public School Students |
| Share who would matriculate at a different college in their choice set | 0.320 | 0.411 | 0.227 | 0.209 | 0.169 | 0.149 |

Table 9
Do Aid Variables that "Should Not Matter" Affect College Choice?
Estimated Odds Ratios from Conditional Logit Regressions
(table and notes continue on next page)

| | I. | II. | III. | IV. |
|---|---|---|---|---|
| Grant (in thousands), specific to the college | **1.110**<br>**(15.890)** | 0.971<br>(-1.310) | **0.937**<br>**(-2.080)** | **0.936**<br>**(-2.130)** |
| Loan (in thousands) from the college | **1.098**<br>**(5.770)** | **1.098**<br>**(5.700)** | **1.110**<br>**(3.880)** | **1.111**<br>**(3.900)** |
| Work study amount (in thousands) from the college | **1.414**<br>**(5.460)** | **1.538**<br>**(6.560)** | **1.548**<br>**(4.900)** | **1.546**<br>**(4.880)** |
| College's tuition (in thousands), in-state or out-of-state as appropriate to the student | **0.980**<br>**(-3.230)** | 0.993<br>(-1.040) | 0.983<br>(-1.650) | 0.984<br>(-1.520) |
| College's room and board (in thousands) | **0.909**<br>**(-3.190)** | 0.947<br>(-1.650) | 0.998<br>(-0.040) | 1.004<br>(0.090) |
| College's per-pupil instructional spending (in thousands) | | | | 0.997<br>(-0.780) |
| Grant is this share of college's per-pupil instructional spending | | | | 0.962<br>(-0.260) |
| Grant is called a "scholarship" | | **1.632**<br>**(4.950)** | **1.394**<br>**(2.540)** | **1.393**<br>**(2.530)** |
| Grant is front-loaded (more in freshman year) | | **1.645**<br>**(2.440)** | **1.930**<br>**(2.230)** | **1.922**<br>**(2.210)** |
| Grant is this share of college's tuition | | 0.794<br>(-0.960) | 0.736<br>(-0.990) | 0.723<br>(-1.030) |
| Grant is this share of college's comprehensive cost | | **30.734**<br>**(3.840)** | **46.146**<br>**(4.020)** | **70.432**<br>**(3.870)** |
| Student's SAT score is this number of percentiles *above* college's average SAT score | **0.951**<br>**(-7.940)** | **0.944**<br>**(-8.690)** | **0.946**<br>**(-6.080)** | **0.945**<br>**(-6.040)** |
| Student's SAT score is this number of percentiles *below* college's average SAT score | 1.003<br>(0.280) | 1.000<br>(0.000) | 0.987<br>(-0.900) | 0.987<br>(-0.880) |
| College is *most* selective to which student was admitted | **1.762**<br>**(8.730)** | **1.757**<br>**(8.560)** | **2.103**<br>**(9.340)** | **2.116**<br>**(9.360)** |
| College is *least* selective to which student was admitted | **0.701**<br>**(-4.150)** | **0.686**<br>**(-4.310)** | **0.786**<br>**(-2.110)** | **0.789**<br>**(-2.080)** |
| Father is alumnus of this college | **1.715**<br>**(3.680)** | **1.655**<br>**(3.330)** | 1.382<br>(1.500) | 1.389<br>(1.520) |
| Mother is alumna of this college | 0.970<br>(-0.160) | 0.978<br>(-0.110) | 1.093<br>(0.310) | 1.085<br>(0.280) |
| Sibling attended or attends this college | **1.949**<br>**(5.280)** | **1.962**<br>**(5.230)** | **1.662**<br>**(2.720)** | **1.674**<br>**(2.750)** |
| Distance between college and student's high school, in hundreds of miles | 1.000<br>(-0.010) | 1.002<br>(0.240) | 0.988<br>(-1.350) | 0.988<br>(-1.340) |
| Square of distance between college and student's high school, in 10000s of miles | 1.000<br>(1.060) | 1.000<br>(1.220) | 1.000<br>(1.430) | 1.000<br>(1.430) |
| Cube of distance between college and student's high school, in 1000000s of miles | 1.000<br>(-1.050) | 1.000<br>(-1.210) | 1.000<br>(-1.390) | 1.000<br>(-1.390) |
| College is in-state for the student | 1.139<br>(1.410) | 1.140<br>(1.400) | 1.025<br>(0.210) | 1.024<br>(0.200) |

| | I. | II. | III. | IV. |
|---|---|---|---|---|
| College is public | 1.096<br>(0.640) | **1.387**<br>**(2.180)** | **1.840**<br>**(2.400)** | **1.860**<br>**(2.430)** |
| Number of observations | 9113 | 9036 | 5455 | 5455 |
| Likelihood ratio ($chi^2$) | 1098.36 | 1202.96 | 770.42 | 771.04 |
| Prob>$chi^2$ | 0 | 0 | 0 | 0 |
| LogLikelihood | -2372.283 | -2296.435 | -1416.548 | -1416.242 |
| Pseudo $R^2$ | 0.188 | 0.208 | 0.214 | 0.214 |

Notes: The table shows results from conditional logit estimation of how a student chooses his matriculation college among the colleges to which he was admitted. The results are shown as odds ratios, with Z-statistics in parentheses below the odd ratios. Results shown in bold print are associated with coefficients that are statistically significantly different from zero with at least 95 percent confidence. The source of data is the College Admissions Project.

Table 10
Students who Gain and Lose Lifetime Present Value by Responding to Aid
Top Panel: Students May Re-Choose Among All Colleges When All Aid is Zeroed Out

| | percentage of students who are in this group | average value of aid | average change in lifetime present value, due to student's response to aid | average change in college consumption, due to student's response to aid | average change in instructional spending, due to student's response to aid | average change in median peer's SAT percentile, due to student's response to aid |
|---|---|---|---|---|---|---|
| students who gain lifetime present value, due to their response to aid | 34.9 | 10,748 | 48,923 | -283 | 2,592 | -3.5 |
| students who lose lifetime present value, due to their response to aid | 31.7 | 2,754 | -81,081 | -1,027 | -16,979 | -8.8 |

The remaining 33.3 percent of students would not change their college choice in response to aid being zeroed out, largely because they actually received little or no aid.

Bottom Panel: Students May Re-Choose Only Among Colleges in their Initial Choice Set When All Aid is Zeroed Out

| | percentage of students who are in this group | average value of aid | average change in lifetime present value, due to student's response to aid | average change in college consumption, due to student's response to aid | average change in instructional spending, due to student's response to aid | average change in median peer's SAT percentile, due to student's response to aid |
|---|---|---|---|---|---|---|
| students who gain lifetime present value, due to their response to aid | 37.2 | 10,798 | 42,685 | -305 | 867 | -3.9 |
| students who lose lifetime present value, due to their response to aid | 29.1 | 2,122 | -43,318 | -1,072 | -9,378 | -8.7 |

The remaining 33.7 percent of students would not change their college choice in response to aid, largely because they actually received little or no aid.

The top and bottom panels of this table are based on predictions from conditional logit equations like that estimated in column II of Table 9. See text for details.

Appendix Table 1
Description of the Colleges to Which Students *Applied*,
from the College Admission Project Data

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| Matriculated at this college | 0.1813 | 0.3853 | 0 | 1 | 17850 |
| Admitted to this college | 0.6566 | 0.4748 | 0 | 1 | 17465 |
| Applied early to this college | 0.1281 | 0.3389 | 0 | 2 | 17694 |
| Withdrew application from this college, usually after early decision elsewhere | 0.0516 | 0.2212 | 0 | 1 | 17694 |
| Grants specific to this college | 1777.8140 | 4933.3550 | 0 | 36000 | 17871 |
| Loans from this college | 413.4718 | 1855.6370 | 0 | 36548 | 17871 |
| Work study amount from this college | 110.7380 | 482.5519 | 0 | 15000 | 17871 |
| Grant is called a "scholarship" | 0.1291 | 0.3354 | 0 | 1 | 17871 |
| Grant is front-loaded (more in freshman year) | 0.0137 | 0.1161 | 0 | 1 | 17871 |
| Grant is this share of tuition | 0.1229 | 0.3676 | 0 | 8 | 17770 |
| Grant is this share of comprehensive cost | 0.0722 | 0.1902 | 0 | 2 | 17758 |
| Student was a recruited athlete at this college | 0.0327 | 0.1779 | 0 | 1 | 17871 |
| Father is an alumnus of this college | 0.0314 | 0.1744 | 0 | 1 | 17871 |
| Mother is an alumna of this college | 0.0209 | 0.1431 | 0 | 1 | 17871 |
| Sibling attended or attends this college | 0.0388 | 0.1932 | 0 | 1 | 17871 |
| College is public | 0.2631 | 0.4403 | 0 | 1 | 17871 |
| College is private not-for-profit | 0.7328 | 0.4436 | 0 | 1 | 17871 |
| College is international, except for Canadian colleges which are treated as U.S. colleges | 0.0040 | 0.0633 | 0 | 1 | 17871 |
| College's median SAT score, in national percentiles | 83.8816 | 12.0390 | 14 | 98 | 16915 |
| Student's SAT score is this many percentiles *above* college's median SAT score | 8.7393 | 9.5927 | 0 | 82 | 16720 |
| Student's SAT score is this many percentiles *below* college's median SAT score | 1.7454 | 5.6654 | 0 | 68 | 16720 |
| In-state tuition | 18181.2300 | 9198.9780 | 0 | 27472 | 17826 |
| Out-of-state tuition | 20497.7600 | 5890.7530 | 0 | 27472 | 17826 |
| Tuition that applies to this student | 19276.9000 | 7965.1400 | 0 | 27472 | 17826 |
| Room and board at this college | 6975.7190 | 1244.3320 | 0 | 10299 | 17796 |
| In-state comprehensive cost of this college | 25745.7900 | 9935.6770 | 0 | 35125 | 17796 |
| Out-of-state comprehensive cost of this college | 28059.7200 | 6681.4230 | 0 | 35125 | 17796 |
| Comprehensive cost that applies to this student | 26841.9800 | 8662.0230 | 0 | 35125 | 17796 |
| Total per-pupil expenditure of this college, in thousands | 68.4888 | 66.0989 | 5 | 816 | 13609 |
| Educational and general per-pupil expenditure of this college, in thousands | 48.2225 | 26.2227 | 5 | 168 | 13609 |
| Instructional per-pupil expenditure of this college, in thousands | 16.8792 | 10.4596 | 2 | 51 | 13609 |
| College is in-state | 0.2666 | 0.4422 | 0 | 1 | 17871 |
| Distance between student's high school and this college, in miles | 673.2152 | 873.1788 | 0 | 5774 | 17146 |
| College is in AK | 0.0001 | 0.0106 | 0 | 1 | 17770 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in AL | 0.0038 | 0.0613 | 0 | 1 | 17770 |
| College is in AR | 0.0003 | 0.0168 | 0 | 1 | 17770 |
| College is in AZ | 0.0039 | 0.0622 | 0 | 1 | 17770 |
| College is in CA | 0.1388 | 0.3458 | 0 | 1 | 17770 |
| College is in CO | 0.0078 | 0.0881 | 0 | 1 | 17770 |
| College is in CT | 0.0533 | 0.2246 | 0 | 1 | 17770 |
| College is in DC | 0.0260 | 0.1591 | 0 | 1 | 17770 |
| College is in DE | 0.0025 | 0.0497 | 0 | 1 | 17770 |
| College is in FL | 0.0111 | 0.1047 | 0 | 1 | 17770 |
| College is in GA | 0.0169 | 0.1290 | 0 | 1 | 17770 |
| College is in HI | 0.0024 | 0.0491 | 0 | 1 | 17770 |
| College is in IA | 0.0032 | 0.0561 | 0 | 1 | 17770 |
| College is in ID | 0.0009 | 0.0300 | 0 | 1 | 17770 |
| College is in IL | 0.0458 | 0.2090 | 0 | 1 | 17770 |
| College is in IN | 0.0166 | 0.1278 | 0 | 1 | 17770 |
| College is in KS | 0.0014 | 0.0375 | 0 | 1 | 17770 |
| College is in KY | 0.0005 | 0.0212 | 0 | 1 | 17770 |
| College is in LA | 0.0070 | 0.0836 | 0 | 1 | 17770 |
| College is in MA | 0.1339 | 0.3406 | 0 | 1 | 17770 |
| College is in MD | 0.0199 | 0.1395 | 0 | 1 | 17770 |
| College is in ME | 0.0159 | 0.1250 | 0 | 1 | 17770 |
| College is in MI | 0.0173 | 0.1303 | 0 | 1 | 17770 |
| College is in MN | 0.0075 | 0.0865 | 0 | 1 | 17770 |
| College is in MO | 0.0217 | 0.1456 | 0 | 1 | 17770 |
| College is in MS | 0.0007 | 0.0260 | 0 | 1 | 17770 |
| College is in MT | 0.0006 | 0.0249 | 0 | 1 | 17770 |
| College is in NC | 0.0411 | 0.1986 | 0 | 1 | 17770 |
| College is in NE | 0.0012 | 0.0344 | 0 | 1 | 17770 |
| College is in NH | 0.0170 | 0.1293 | 0 | 1 | 17770 |
| College is in NJ | 0.0311 | 0.1735 | 0 | 1 | 17770 |
| College is in NM | 0.0011 | 0.0327 | 0 | 1 | 17770 |
| College is in NV | 0.0005 | 0.0225 | 0 | 1 | 17770 |
| College is in NY | 0.1187 | 0.3235 | 0 | 1 | 17770 |
| College is in OH | 0.0201 | 0.1405 | 0 | 1 | 17770 |
| College is in OK | 0.0011 | 0.0335 | 0 | 1 | 17770 |
| College is in OR | 0.0058 | 0.0759 | 0 | 1 | 17770 |
| College is in PA | 0.0723 | 0.2589 | 0 | 1 | 17770 |
| College is in RI | 0.0320 | 0.1761 | 0 | 1 | 17770 |
| College is in SC | 0.0037 | 0.0604 | 0 | 1 | 17770 |
| College is in TN | 0.0106 | 0.1023 | 0 | 1 | 17770 |
| College is in TX | 0.0185 | 0.1346 | 0 | 1 | 17770 |
| College is in UT | 0.0032 | 0.0565 | 0 | 1 | 17770 |
| College is in VA | 0.0361 | 0.1866 | 0 | 1 | 17770 |
| College is in VT | 0.0110 | 0.1042 | 0 | 1 | 17770 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in WA | 0.0088 | 0.0936 | 0 | 1 | 17770 |
| College is in WI | 0.0061 | 0.0781 | 0 | 1 | 17770 |
| College is in WV | 0.0001 | 0.0075 | 0 | 1 | 17770 |
| College is in WY | 0.0003 | 0.0168 | 0 | 1 | 17770 |

Appendix Table 2
Description of the Colleges at Which Students *Matriculated*,
from the College Admission Project Data

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| Matriculated at this college | 1.0000 | 0.0000 | 1 | 1 | 3236 |
| Admitted to this college | 1.0000 | 0.0000 | 1 | 1 | 3236 |
| Applied early to this college | 0.3142 | 0.4722 | 0 | 2 | 3221 |
| Withdrew application from this college, usually after early decision elsewhere | 0.0000 | 0.0000 | 0 | 0 | 3221 |
| Grants specific to this college | 4029.0040 | 7051.1670 | 0 | 36000 | 3236 |
| Loans from this college | 1020.0040 | 2721.6190 | 0 | 36348 | 3236 |
| Work study amount from this college | 296.3472 | 768.4207 | 0 | 15000 | 3236 |
| Grant is called a "scholarship" | 0.2692 | 0.4436 | 0 | 1 | 3236 |
| Grant is front-loaded (more in freshman year) | 0.0343 | 0.1820 | 0 | 1 | 3236 |
| Grant is this share of tuition | 0.2875 | 0.5517 | 0 | 7 | 3201 |
| Grant is this share of comprehensive cost | 0.1665 | 0.2728 | 0 | 2 | 3196 |
| Student was a recruited athlete at this college | 0.0402 | 0.1964 | 0 | 1 | 3236 |
| Father is an alumnus of this college | 0.0664 | 0.2491 | 0 | 1 | 3236 |
| Mother is an alumna of this college | 0.0396 | 0.1949 | 0 | 1 | 3236 |
| Sibling attended or attends this college | 0.0831 | 0.2761 | 0 | 1 | 3236 |
| College is public | 0.2843 | 0.4512 | 0 | 1 | 3236 |
| College is private not-for-profit | 0.7086 | 0.4562 | 0 | 1 | 3236 |
| College is international, except for Canadian colleges which are treated as U.S. colleges | 0.0068 | 0.0822 | 0 | 1 | 3236 |
| College's median SAT score, in national percentiles | 83.4215 | 12.5494 | 32 | 98 | 3033 |
| Student's SAT score is this many percentiles *above* college's median SAT score | 8.4548 | 9.1831 | 0 | 53 | 2990 |
| Student's SAT score is this many percentiles *below* college's median SAT score | 1.4351 | 4.8994 | 0 | 50 | 2990 |
| In-state tuition | 17431.8300 | 9512.6270 | 0 | 27472 | 3218 |
| Out-of-state tuition | 19841.1300 | 6370.6670 | 0 | 27472 | 3218 |
| Tuition that applies to this student | 18340.3700 | 8599.1560 | 0 | 27472 | 3218 |
| Room and board at this college | 6821.8120 | 1352.4620 | 0 | 10299 | 3209 |
| In-state comprehensive cost of this college | 24881.0900 | 10409.1500 | 0 | 35125 | 3209 |
| Out-of-state comprehensive cost of this college | 27285.9500 | 7335.3150 | 0 | 35125 | 3209 |
| Comprehensive cost that applies to this student | 25792.1800 | 9469.9140 | 0 | 35125 | 3209 |
| Total per-pupil expenditure of this college, in thousands | 67.5188 | 64.2833 | 5 | 816 | 2434 |
| Educational and general per-pupil expenditure of this college, in thousands | 47.9875 | 26.4929 | 5 | 168 | 2434 |
| Instructional per-pupil expenditure of this college, in thousands | 16.6971 | 10.2716 | 2 | 51 | 2434 |
| College is in-state | 0.3368 | 0.4727 | 0 | 1 | 3236 |
| Distance between student's high school and this college, in miles | 575.6313 | 827.2526 | 0 | 5769 | 3094 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in AK | 0.0000 | 0.0000 | 0 | 0 | 3203 |
| College is in AL | 0.0050 | 0.0705 | 0 | 1 | 3203 |
| College is in AR | 0.0006 | 0.0250 | 0 | 1 | 3203 |
| College is in AZ | 0.0053 | 0.0727 | 0 | 1 | 3203 |
| College is in CA | 0.1199 | 0.3249 | 0 | 1 | 3203 |
| College is in CO | 0.0094 | 0.0963 | 0 | 1 | 3203 |
| College is in CT | 0.0537 | 0.2255 | 0 | 1 | 3203 |
| College is in DC | 0.0265 | 0.1608 | 0 | 1 | 3203 |
| College is in DE | 0.0022 | 0.0467 | 0 | 1 | 3203 |
| College is in FL | 0.0203 | 0.1410 | 0 | 1 | 3203 |
| College is in GA | 0.0131 | 0.1138 | 0 | 1 | 3203 |
| College is in HI | 0.0044 | 0.0660 | 0 | 1 | 3203 |
| College is in IA | 0.0025 | 0.0499 | 0 | 1 | 3203 |
| College is in ID | 0.0022 | 0.0467 | 0 | 1 | 3203 |
| College is in IL | 0.0571 | 0.2321 | 0 | 1 | 3203 |
| College is in IN | 0.0190 | 0.1367 | 0 | 1 | 3203 |
| College is in KS | 0.0025 | 0.0499 | 0 | 1 | 3203 |
| College is in KY | 0.0006 | 0.0250 | 0 | 1 | 3203 |
| College is in LA | 0.0050 | 0.0705 | 0 | 1 | 3203 |
| College is in MA | 0.1218 | 0.3271 | 0 | 1 | 3203 |
| College is in MD | 0.0187 | 0.1356 | 0 | 1 | 3203 |
| College is in ME | 0.0140 | 0.1177 | 0 | 1 | 3203 |
| College is in MI | 0.0194 | 0.1378 | 0 | 1 | 3203 |
| College is in MN | 0.0053 | 0.0727 | 0 | 1 | 3203 |
| College is in MO | 0.0212 | 0.1442 | 0 | 1 | 3203 |
| College is in MS | 0.0012 | 0.0353 | 0 | 1 | 3203 |
| College is in MT | 0.0012 | 0.0353 | 0 | 1 | 3203 |
| College is in NC | 0.0390 | 0.1937 | 0 | 1 | 3203 |
| College is in NE | 0.0022 | 0.0467 | 0 | 1 | 3203 |
| College is in NH | 0.0172 | 0.1299 | 0 | 1 | 3203 |
| College is in NJ | 0.0284 | 0.1662 | 0 | 1 | 3203 |
| College is in NM | 0.0009 | 0.0306 | 0 | 1 | 3203 |
| College is in NV | 0.0022 | 0.0467 | 0 | 1 | 3203 |
| College is in NY | 0.1065 | 0.3085 | 0 | 1 | 3203 |
| College is in OH | 0.0178 | 0.1322 | 0 | 1 | 3203 |
| College is in OK | 0.0022 | 0.0467 | 0 | 1 | 3203 |
| College is in OR | 0.0078 | 0.0880 | 0 | 1 | 3203 |
| College is in PA | 0.0743 | 0.2623 | 0 | 1 | 3203 |
| College is in RI | 0.0300 | 0.1705 | 0 | 1 | 3203 |
| College is in SC | 0.0066 | 0.0807 | 0 | 1 | 3203 |
| College is in TN | 0.0140 | 0.1177 | 0 | 1 | 3203 |
| College is in TX | 0.0225 | 0.1483 | 0 | 1 | 3203 |
| College is in UT | 0.0091 | 0.0947 | 0 | 1 | 3203 |
| College is in VA | 0.0406 | 0.1974 | 0 | 1 | 3203 |

| Variable | Mean | Std. Dev. | Minimum | Maximum | No. of Obs. |
|---|---|---|---|---|---|
| College is in VT | 0.0106 | 0.1025 | 0 | 1 | 3203 |
| College is in WA | 0.0094 | 0.0963 | 0 | 1 | 3203 |
| College is in WI | 0.0059 | 0.0768 | 0 | 1 | 3203 |
| College is in WV | 0.0000 | 0.0000 | 0 | 0 | 3203 |
| College is in WY | 0.0006 | 0.0250 | 0 | 1 | 3203 |

Appendix Table 3
Are Students from Low and High Income Families Equally Sensitive to Aid Variables that "Should Not Matter"?
*Selected* Estimated Odds Ratios from Conditional Logit Regressions

| | Parents Income is: | | | | same effect for all groups? |
|---|---|---|---|---|---|
| | Low | Medium Low | Medium High | High | |
| Grant is called a "scholarship" | 2.213 (2.670) | 2.291 (3.990) | 1.842 (3.680) | 1.194 (0.800) | rejected |
| Grant is front-loaded (more in freshman year) | 1.256 (0.320) | 1.142 (0.300) | 2.139 (2.250) | 1.795 (1.420) | |
| Grant is this share of college's tuition | 0.749 (-0.430) | 0.562 (-1.340) | 1.074 (0.170) | 0.563 (-0.900) | |
| Grant is this share of college's comprehensive cost | 23.613 (1.280) | 37.709 (2.790) | 9.634 (1.510) | 8.033 (1.430) | rejected |
| regression includes all other variables shown in Table 9? | yes | yes | yes | yes | |

Appendix Table 4
Are Students whose Parents Attended More and Less Selective Colleges Equally Sensitive to Aid Variables that "Should Not Matter"?
*Selected* Estimated Odds Ratios from Conditional Logit Regressions

| | Parents' College is: | | | same effect for all groups? |
|---|---|---|---|---|
| | low selectivity | medium selectivity | high selectivity | |
| Grant is called a "scholarship" | 1.702 (4.390) | 1.703 (2.450) | 1.173 (0.470) | rejected |
| Grant is front-loaded (more in freshman year) | 1.438 (1.510) | 1.805 (1.130) | 2.659 (1.500) | |
| Grant is this share of college's tuition | 0.770 (-0.870) | 0.691 (-0.830) | 2.869 (0.760) | |
| Grant is this share of college's comprehensive cost | 35.634 (3.300) | 41.044 (2.040) | 0.729 (-0.070) | rejected |
| regression includes all other variables shown in Table 9? | yes | yes | yes | |

Appendix Table 5
Are Students from Public and Private Schools Equally Sensitive to Aid Variables that "Should Not Matter"?
*Selected* Estimated Odds Ratios from Conditional Logit Regressions

| | Public High School | Private High School | same effect for all groups? |
|---|---|---|---|
| Grant is called a "scholarship" | 1.685 (4.080) | 1.638 (3.040) | |
| Grant is front-loaded (more in freshman year) | 1.661 (2.140) | 1.300 (0.620) | |
| Grant is this share of college's tuition | 0.774 (-0.910) | 0.730 (-0.600) | |
| Grant is this share of college's comprehensive cost | 31.738 (3.130) | 34.150 (2.180) | |
| regression includes all other variables shown in Table 9? | yes | yes | |