This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: College Choices: The Economics of Where to Go, When to Go, and How to Pay For It

Volume Author/Editor: Caroline M. Hoxby, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-35535-7

Volume URL: http://www.nber.org/books/hoxb04-1

Conference Date: August 13-15, 2002

Publication Date: September 2004

Title: Do and Should Financial Aid Packages Affect Students' College Choices?

Author: Christopher Avery, Caroline Minter Hoxby

URL: http://www.nber.org/chapters/c10102

6

Do and Should Financial Aid Packages Affect Students' College Choices?

Christopher Avery and Caroline M. Hoxby

6.1 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 usable only at in-state

public colleges; state scholarships that are usable at any in-state college; work-study programs; 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.

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 not a large group, but they are important. 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 students is also important purely for reasons of scientific inquiry. 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 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. Finally, high-aptitude students are likely to be the investors who most closely obey the model of the 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. 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. See, for instance, Ehrenberg and Sherman's (1984) study of students who were accepted by Cornell University in the spring of 1981. While we believe that

college's internal analyses provide helpful evidence, they have flaws: They necessarily focus on a narrow set of students (the students accepted by one college); they are sometimes tacit about their empirical methods; and the studies are hard to compare because most are unpublished (distribution is often purely internal). Also, while colleges have complete information about their own aid offers and matriculation, they typically have poor information about their admittees' other college acceptances and aid offers.

Though there is a 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 Dynarski (chap. 2), Long (chap. 3), and Bettinger (chap. 5) in this volume.

6.2   How *Should* Students Respond to Scholarships and Aid?

6.2.1   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 in the following), then he need make only two calculations for each college in his choice set. Supposing that the

---

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.

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:

(1)
$$\sum_{t=1}^{t=4} \frac{(\text{TuitionFees}_{jt} + \text{RoomBoard}_{jt})}{(1 + \delta)^{t-1}}$$

$$-\sum_{t=1}^{t=4} \frac{\text{ApplicableGrants}_{ijt} + \text{ApplicableLoans}^*_{ijt} + \text{WorkStudySubsidies}^*_{ijt})}{(1 + \delta)^{t-1}}$$

$$+\sum_{t=1}^{t=T} \frac{\text{LoanRepayments}^*_{ijt}}{(1 + \delta)^{t-1}}.$$

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

(2)
$$\sum_{t=1}^{t=4} \cdot$$

$$\frac{\text{FoodConsumption}_{ijt} + \text{HousingConsumption}_{ijt} + \text{OtherConsumption}_{ijt}}{(1 + \delta)^{t-1}}$$

$$\cdot \sum_{t=5}^{t=T_i} \frac{r_{it}(\text{Resources}_{jt} + \text{PeerSpillovers}_{jt})}{(1 + \delta)^{t-1}}.$$

In both equations (1) and (2), i indexes individual students, j indexes colleges, δ is the discount rate on future years, and t indexes years ($t = 1$ is the freshman year, $t = 5$ is the first postbaccalaureate year, and T_i is the end of working 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 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 working 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]

---

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

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, etc.) 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.[3] Notice that we allow the student's return on the stock of human capital he acquires in college to be specific to him (because of ability) and specific to each year (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 nonfinancial returns.

The alert reader may notice that we have said nothing about the opportunity cost of college, which is 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 choose which college to attend, the student has only to subtract equation (1) from equation (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.

---

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 less-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 differential associated with the 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.

3. Good peers may merely facilitate a student's absorbing the nonpeer resources invested in him; in this case the functional form, though not the spirit, of equation (2) should be altered. Specifically, the equation should include terms that interact with peer quality and resources, not the level of peer quality.

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 group of students 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 peers, 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 students 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 oversubscribed 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. As a consequence, college C's peer quality would no longer be equivalent to that of colleges A and B. Of course, it is possible that an *individual* student will face no trade-off between two colleges in his choice set. However, such no-trade-off situations must necessarily be idiosyncratic to individual students. They cannot hold generally.

We will 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 will 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 will 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 will act in accordance with the present discounted value of various forms of aid—for instance, he will recognize 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 last few sentences. (The word "rational" can be loaded with meaning about mental processing. We are using it in a strictly limited way. In this chapter, "rational" means that a student obeys the standard model of human capital investment.)

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 not attempt to maximize his own lifetime utility when he chooses a college.

6.3   The College Admissions Project Survey Data

Our data come from the College Admissions Project, in which we surveyed high school seniors applying to college during the 1999–2000 academic year.[4] 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 in typical surveys that their behavior cannot be analyzed, even if the survey 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 hope to learn how students who can attract interesting aid packages respond to them.

6.3.1   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) and the experience of admissions experts (Andrew Fairbanks, Michael Behnke, and Larry Momo). 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 percent of the senior class, while counselors at private schools (which tend to be smaller and have higher mean college aptitude) selected students at random from the top 20 percent of the senior class.[5] The counselors distributed the surveys to students, collected the completed surveys, and returned them to us for coding.[6] 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

---

4. See Avery and Hoxby (2000) for additional detail.

5. The counselors were given detailed instructions for random sampling from the top twenty, thirty, forty, or fifty students in the senior class, depending on the size of the school. For example, a counselor from a public school with 200 students in a class 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.

6. 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.

academic year. The first questionnaire was administered in January 2000. It asked for the same background, academic, and extracurricular information that college applications require. The majority of these questions were taken directly from the Common Application, which is accepted by many colleges in place of their specific application forms. In addition, each student listed (up to) his ten most preferred colleges regardless of whether he had applied to them yet. Each student also 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).[7] The responses on merit-based scholarships, both institutional and outside, were accurate and clear, presumably because students were proud of them as accomplishments.[8] 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 6.1 for the income categories.)

We matched the College Admissions Project data to colleges' 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 2000–2001 school year, which corresponds to the survey participants' freshman year.[9]

The College Admissions Project survey produced a response rate of approximately 65 percent, including information for 3,240 students from 396 high schools. So far as we could discern from the data we had on respon-

7. 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 minority of students did so.

8. In most cases, we were able to validate the terms of the scholarship because it is described on the college's website or in its publications. Our survey respondents' descriptions were very accurate; in no case did we fail to validate the key terms of a scholarship.

9. We collected the administrative data from the following sources in order: 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; U.S. Department of Education 2001, 2002), the College Board's annual survey (ACS; 2002), the 2001 edition of *Peterson's Guide to Four-Year Colleges* (2002), and the 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.

Table 6.1   Description of the Students in the College Admission Project Data

Variable	Mean	Standard Deviation	Minimum	Maximum
Male	0.4120	0.4923	0	1
White non-Hispanic	0.7321	0.4429	0	1
Black	0.0350	0.1837	0	1
Asian	0.1571	0.3640	0	1
Hispanic	0.0382	0.1918	0	1
Native American	0.0010	0.0313	0	1
Other race/ethnicity	0.0366	0.1878	0	1
Parents are married	0.8305	0.3752	0	1
Sibling(s) enrolled in college	0.2327	0.4226	0	1
Parents' income, estimated if necessary	119,929.0000	65,518.2100	9,186	240,000
Parents' income < $20k	0.0221	0.1469	0	1
Parents' income $20–30k	0.0379	0.1910	0	1
Parents' income $30–40k	0.0301	0.1710	0	1
Parents' income $40–50k	0.0398	0.1955	0	1
Parents' income $50–60k	0.0497	0.2174	0	1
Parents' income $60–70k	0.0594	0.2363	0	1
Parents' income $70–80k	0.0690	0.2535	0	1
Parents' income $80–90k	0.0522	0.2225	0	1
Parents' income $90–100k	0.0855	0.2796	0	1
Parents' income $100–120k	0.1495	0.3566	0	1
Parents' income $120–140k	0.0923	0.2895	0	1
Parents' income $140–160k	0.0771	0.2667	0	1
Parents' income $160–200k	0.0761	0.2653	0	1
Parents' income $200+k	0.1594	0.3661	0	1
Expected family contribution, estimated if necessary	27,653.4700	16,523.9200	0	120,000
Applied for financial aid?	0.5946	0.4910	0	1
Finances influenced college choice?	0.4114	0.4922	0	1
Amount of outside scholarships, applicable at *any* college	203.0781	799.9640	0	12,500
National Merit Scholarship winner	0.0494	0.2167	0	1
Student's SAT score, sum of math and verbal, converted from ACT score if necessary	1,356.9110	138.8193	780	1,600
Student's SAT score, expressed as national percentile	90.4013	12.3362	12	100
Median SAT score at *most* selective college to which student was admitted	86.4092	10.3836	34	98
Median SAT score at *least* selective college to which student was admitted	73.8469	14.5646	14	97
Number of colleges to which student was admitted	3.5250	2.1293	1	10
Student's high school was private	0.4534	0.4979	0	1
Student's high school in AL	0.0170	0.1292	0	1
Student's high school in AR	0.0028	0.0526	0	1

(*continued*)

Table 6.1 (continued)

Variable	Mean	Standard Deviation	Minimum	Maximum
Student's high school in AZ	0.0093	0.0958	0	1
Student's high school in CA	0.1222	0.3276	0	1
Student's high school in CO	0.0120	0.1091	0	1
Student's high school in CT	0.0327	0.1779	0	1
Student's high school in DC	0.0096	0.0974	0	1
Student's high school in FL	0.0287	0.1670	0	1
Student's high school in GA	0.0111	0.1048	0	1
Student's high school in HI	0.0201	0.1402	0	1
Student's high school in ID	0.0031	0.0555	0	1
Student's high school in IL	0.0633	0.2435	0	1
Student's high school in IN	0.0086	0.0926	0	1
Student's high school in KS	0.0046	0.0679	0	1
Student's high school in KY	0.0031	0.0555	0	1
Student's high school in LA	0.0105	0.1019	0	1
Student's high school in MA	0.0855	0.2797	0	1
Student's high school in MD	0.0327	0.1779	0	1
Student's high school in ME	0.0052	0.0723	0	1
Student's high school in MI	0.0198	0.1392	0	1
Student's high school in MN	0.0056	0.0743	0	1
Student's high school in MO	0.0198	0.1392	0	1
Student's high school in MT	0.0019	0.0430	0	1
Student's high school in NC	0.0219	0.1464	0	1
Student's high school in NE	0.0031	0.0555	0	1
Student's high school in NH	0.0167	0.1280	0	1
Student's high school in NJ	0.0522	0.2224	0	1
Student's high school in NM	0.0102	0.1004	0	1
Student's high school in NV	0.0031	0.0555	0	1
Student's high school in NY	0.1278	0.3339	0	1
Student's high school in OH	0.0309	0.1730	0	1
Student's high school in OK	0.0062	0.0783	0	1
Student's high school in OR	0.0105	0.1019	0	1
Student's high school in PA	0.0472	0.2121	0	1
Student's high school in RI	0.0086	0.0926	0	1
Student's high school in SC	0.0031	0.0555	0	1
Student's high school in TN	0.0201	0.1402	0	1
Student's high school in TX	0.0395	0.1948	0	1
Student's high school in UT	0.0071	0.0840	0	1
Student's high school in VA	0.0333	0.1795	0	1
Student's high school in VT	0.0031	0.0555	0	1
Student's high school in WA	0.0160	0.1257	0	1
Student's high school in WI	0.0077	0.0875	0	1
Student's high school in WU	0.0028	0.0526	0	1

Source: 3,240 students in College Admissions Project sample.

dents, partial respondents, and nonrespondents, lack of participation was uncorrelated with observable student and school characteristics. This was probably because counselor vagaries accounted for most of the partial responses and nonresponses.[10] The final sample contains students from forty-three states plus the District of Columbia.[11] Although the sample was constructed to include students from every region of the country, it is intentionally representative of students who apply to highly selective colleges and, therefore, nonrepresentative 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.

6.3.2 The Typical Student in the College Admissions Project

The summary statistics shown in tables 6.1 and 6.2 (and in tables 6A.2 and 6A.3) demonstrate that the sample is quite special. The average (combined verbal and math) Scholastic Aptitude Test (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. Forty-five percent of the students attended private school, and their parents' income averaged $119,929 in 1999. However, 76 percent of the sample had incomes below the cutoff where a family is considered for aid by selective private colleges (the cutoff averaged $160,000 for 2000–2001, but the actual cutoff depended on family circumstances). Fifty-nine percent of the students applied for need-based financial aid, and 41 percent of the families reported that finances influenced their college choice.[12] 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.

Eighty-three percent of the student's parents were currently married,

10. The most common reasons for failure to return the survey were changes in the job of the high school counselor (so that the survey would no longer be a natural part of his job), the counselor's becoming pregnant or ill, and other administrative problems that were unrelated to the college admissions outcomes of students who had been selected to participate. We tested whether respondents, partial respondents, and nonrespondents differed in school characteristics, January survey characteristics, and basic characteristics reported by counselors (sex, race, class rank). We did not find any statistically significant differences among respondents, partial respondents, and nonrespondents.

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

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

Table 6.2   **Description of the Colleges to Which Students Were Admitted, from the College Admission Project Data**

Variable	Mean	Standard Deviation	Minimum	Maximum
Matriculated at this college	0.2825	0.4502	0	1
Admitted to this college	1.0000	0.0000	1	1
Applied early to this college	0.1298	0.3405	0	2
Withdrew application from this college, usually after early decision elsewhere	0.0000	0.000	0	0
Grants specific to this college	2,719.8600	5,870.0240	0	36,000
Loans from this college	641.3459	2,282.1720	0	36,548
Work study amount from this college	172.1048	593.0736	0	15,000
Grant is called a "scholarship"	0.1958	0.3968	0	1
Grant is front-loaded (more in freshman year)	0.0212	0.1440	0	1
Grant is this share of tuition	0.1885	0.4369	0	7
Grant is this share of comprehensive cost	0.1109	0.2258	0	2
Student was a recruited athlete at this college	0.0275	0.1634	0	1
Father is an alumnus of this college	0.0401	0.1962	0	1
Mother is an alumna of this college	0.0283	0.1659	0	1
Sibling attended or attends this college	0.0484	0.2146	0	1
College is public	0.3325	0.4711	0	1
College is private not-for-profit	0.6628	0.4737	0	1
College is international, except for Canadian colleges, which are treated as U.S. colleges	0.0045	0.0672	0	1
College's median SAT score, in national percentiles	80.5947	12.5188	14	98
Student's SAT score is this many percentiles *above* college's median SAT score	11.2945	10.2160	0	82
Student's SAT score is this many percentiles *below* college's median SAT score	1.1006	4.3038	0	58
In-state tuition	16,435.1500	9,594.0020	0	27,472
Out-of-state tuition	19,293.5700	6,190.8330	0	27,472
Tuition that applies to this student	17,670.6000	8,491.8630	0	27,472
Room and board at this college	6,808.9370	1,322.2720	0	10,299
In-state comprehensive cost of this college	23,785.2000	10,368.3300	0	35,125
Out-of-state comprehensive cost of this college	26,641.5400	7,032.6210	0	35,125
Comprehensive cost that applies to this student	25,022.2000	9,219.1590	0	35,125
Per-pupil expenditure on students (instruction, student services, academic support, scholarships) of this college, in thousands	26.0321	15.5894	2	146
Instructional per-pupil expenditure of this college, in thousands	17.4502	11.8691	2	72
College is in-state	0.3270	0.4691	0	1
Distance between student's high school and this college, in miles	597.1856	808.9188	0	5,769
College is in AK	0.0000	0.0000	0	0

Table 6.2   (continued)

Variable	Mean	Standard Deviation	Minimum	Maximum
College is in AL	0.0053	0.0724	0	1
College is in AR	0.0004	0.0187	0	1
College is in AZ	0.0056	0.0748	0	1
College is in CA	0.1385	0.3454	0	1
College is in CO	0.0109	0.1038	0	1
College is in CT	0.0380	0.1913	0	1
College is in DC	0.0260	0.1591	0	1
College is in DE	0.0032	0.0561	0	1
College is in FL	0.0164	0.1271	0	1
College is in GA	0.0197	0.1389	0	1
College is in HI	0.0035	0.0592	0	1
College is in IA	0.0042	0.0648	0	1
College is in ID	0.0013	0.0363	0	1
College is in IL	0.0543	0.2265	0	1
College is in IN	0.0206	0.1422	0	1
College is in KS	0.0022	0.0468	0	1
College is in KY	0.0006	0.0248	0	1
College is in LA	0.0094	0.0965	0	1
College is in MA	0.1054	0.3070	0	1
College is in MD	0.0219	0.1462	0	1
College is in ME	0.0144	0.1191	0	1
College is in MI	0.0227	0.1488	0	1
College is in MN	0.0089	0.0938	0	1
College is in MO	0.0259	0.1589	0	1
College is in MS	0.0009	0.0296	0	1
College is in MT	0.0010	0.0311	0	1
College is in NC	0.0356	0.1852	0	1
College is in NE	0.0018	0.0419	0	1
College is in NH	0.0118	0.1078	0	1
College is in NJ	0.0217	0.1457	0	1
College is in NM	0.0017	0.0408	0	1
College is in NV	0.0008	0.0281	0	1
College is in NY	0.1212	0.3263	0	1
College is in OH	0.0273	0.1630	0	1
College is in OK	0.0018	0.0419	0	1
College is in OR	0.0087	0.0928	0	1
College is in PA	0.0713	0.2573	0	1
College is in RI	0.0193	0.1376	0	1
College is in SC	0.0049	0.0700	0	1
College is in TN	0.0139	0.1170	0	1
College is in TX	0.0222	0.1474	0	1
College is in UT	0.0045	0.0668	0	1
College is in VA	0.0391	0.1938	0	1
College is in VT	0.0104	0.1013	0	1
College is in WA	0.0122	0.1098	0	1
College is in WI	0.0090	0.0942	0	1
College is in WV	0.0000	0.0000	0	0
College is in WY	0.0003	0.0162	0	1

Source: 11,468 college admissions events for the 3,240 students in the College Admissions Project sample.

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 non-Hispanic, and 3.8 percent Hispanic. We found that the black and Hispanic subgroups were too small for separate analysis. We also found that the white and Asian subgroups behaved similarly, all else being equal. Thus, we will not discuss students' races further in this chapter.

Looking at table 6A.2, which shows descriptive statistics on the colleges where the students *applied,* 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 6.2 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 6.2 to table 6A.2, 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 a 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. Sixty-six percent of the colleges to which they were admitted were private, and their mean tuition was $17,671. Notice that we show 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, table 6A.3 shows descriptive statistics for 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 fifty 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.

6.3.3   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 crosswalk 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 2000–2001.[13] 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.

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 toward the student's college education.) We can explain 88 percent of the variation in the Expected Family Contribution using just two variables: parents' income and likely current expenditures for the college education of older siblings. We know about siblings' enrollment and likely expenditures for their education. Therefore, our estimates of parents' income based on the Expected Family Contribution and siblings' college expenses are highly accurate. Later, readers will see that we only need to group parents into four income groups. For families that reported both 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 1999 income from his occupation). For families for which we could check this method, we found that it assigned them to the correct income group 91 percent of the time.[14]

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 a four-

13. This is an important, although often neglected, conversion. A given change in an SAT scale score (of, say, 100 points) corresponds to a differing number of percentiles, depending on where the scale score is in the distribution. For instance, the difference between a combined score of 1500 and 1600 is only a few percentiles, but the difference between a combined score of 1400 and 1500 is three to four times as many percentiles. The unconverted scale scores generate seriously biased estimates when used in regression equations where the scores enter linearly.

14. For the occupation-based estimate of parents' income, nearly all of the "mistakes" were caused by our assigning families to the medium-high-income group when they truly belonged in the high-income group. We suspect that our medium-high-income group probably contains about twelve families that should be assigned to the high-income group. The underassignment to the high-income group is caused by professional occupations' having income distributions with a right-hand skew. Think, for instance, of attorneys. Parents make it into the high-income group because they are, say, unusually highly paid lawyers, not because the mean income for a lawyer would put them there.

year total rather than an annual amount, or recorded with insufficient re-strictions. In all cases where a student reported a named scholarship or grant with published parameters (for instance, Morehouse Scholars at the University of North Carolina), we validated the basic terms of the schol-arship or grant.

6.4   Empirical Strategy

Our empirical strategy is straightforward. We are interested in discover-ing 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 picked—in our case, it is a dummy variable equal to 1 for the college at which the stu-dent matriculated and 0 for all of the other colleges in the student's 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 fol-lowing log likelihood function:

$$(3) \qquad \ln L = \sum_{i=1}^{n} \sum_{j=1}^{J_i} \text{matric}_{ij} \ln \text{Prob}(\text{collegechoice}_i = j),$$

where

$$(4) \qquad \text{Prob}(\text{collegechoice}_i = j) = \frac{e^{\boldsymbol{\beta}' \mathbf{x}_{ij}}}{\sum_{j=1}^{J_i i} e^{\boldsymbol{\beta}' \mathbf{x}_{ij}}}.$$

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

In equations (3) and (4), i indexes the student; j indexes the college; the indicator variable matric_{ij} is equal to 1 if student i chooses to matriculate at college j, and zero otherwise; and collegechoice_i is simply the student's col-lege choice. The vector $\mathbf{x}_{ij}$ includes the attributes of choice j for student i—note that the subscripts indicate that the attributes may be match specific. $\boldsymbol{\beta}$ is the vector of effects that we are interested in estimating. One maxi-mizes the log of the likelihood simply to make estimation easier.

The choice problem we are investigating is suitable for conditional logit estimation but unsuitable for multinomial logit estimation. Although multinomial logit is related to conditional logit and sometimes confused

with it, it 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.

6.4.1   The Variation That Drives the Estimates

It is important to be explicit about the variation that drives our estimates. First, all of the variation used is *within* the choice set of a student. Second, within a given student's choice set, there is variation in the attributes of colleges because colleges vary for reasons 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 available to him given his characteristics and aptitude.

There is one possible worry about the endogeneity of the attributes of colleges, 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 only a function of his merit but also reflects effort on the student's part *that is observed by that college only* and that is a function of his desire to attend that particular college. Note our emphasis on the effort being observed by one college only. Any merit or effort that can be observed by all colleges is *not* a problem. For instance, if a student collects special letters or other evidence of merit that he sends only to his most preferred college, it could cause a problem (the aid package offered by his most preferred college might be a function of a match-specific liking). If he collects the special letters or evidence of merit with one college in mind but actually sends them with all his applications, there is no problem.

There are three reasons to think that this form of endogeneity is minor. First, even if a student gathers special information with only one college in mind, he is still best off sending it to all colleges. Second, colleges ask for information in a calculating way. Although a student who supplies unasked-for information may improve his aid package somewhat, he is unlikely to get it changed radically because most important achievements are revealed in the application materials. Third, 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 Admissions 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 very minor in character, based on the verbatim responses of students who described the revision they obtained. Only two students described a revision that would cost a college more than $1,000 over four years.

6.4.2   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 $\mathbf{x}_{ij}$ for conditional logit estimation. However, a little thought shows why this is so. The student's own characteristics are the 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 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, the difference between a student's SAT score and a college's average SAT score is a match-specific attribute included in the vector $\mathbf{x}_{ij}$. College attributes that differ across colleges but are constant across students within a college (such as whether the college is publicly or privately controlled) are also in the vector $\mathbf{x}_{ij}$ because 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 $\mathbf{x}_{ij}$ 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, such as region, gender, and race. We show every dimension for which the data even hinted at there being differential responsiveness.

6.4.3   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. In the appendix, we offer more detail on this issue.[15] Here, we offer the logic of the situation.

Suppose that we *did* observe all of the colleges to which student could

---

15. 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.

have been admitted and the financial aid packages he would have received at each college.[16] Then it would be the case that, given the student's preferences, some colleges were dominated by others. The dominated colleges would be irrelevant to the student's choice. 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 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 revealing that he expects the college to be a relevant alternative in some scenario. Put another way, when a student chooses not to apply to a college, he is revealing that the college will be dominated in all likely scenarios. Dominated colleges may be very similar to (but, nevertheless, less preferred than) other colleges in the student's choice set. By forming his choice set this way (called "endogenous choice set formation"), the student is helping to exclude irrelevant alternatives from his choice set. Irrelevant alternatives pose a problem for conditional logit estimation.

In addition, some readers may be comforted by the fact the data do not reject in the Hausman-type test of the independence of irrelevant alternatives (see appendix).

6.4.4   Early Decision Applicants

When a student applies early decision, he gets the admission benefits of an early decision application (a slight relaxation of the admissions standard) in return for giving up the right to use the knowledge conveyed by the regular admissions process (he cannot bargain with the early decision college using admissions and aid offers from other colleges). Essentially, he predicts what his other alternatives would have been and chooses to apply early decision based on those predictions, his preferences, and his beliefs about the relaxation in standards for early decision applicants.

Because our survey asks about the ten colleges most under consideration by a student, we know which colleges the student considered relevant even if he applies early decision. However, for some early decision applicants, we do not know what admissions and aid outcomes they would have received at the colleges to which they were applying through regular admission. We lack such information when a student is accepted early decision and, consequently, withdraws his regular applications. Some of our early decision applicants withdraw their applications before learning about their alternative admissions outcomes; other early decision applicants report admissions and aid outcomes from regular applications, which suggests that they got this information before withdrawing. (Early *action* applicants

---

16. We are simplifying for the purpose of exposition. We might actually want to estimate the probability distribution of aid packages for each college. 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.

rarely withdraw their regular applications, it being to their advantage to have as wide a menu as possible.) Out of 3,240 students, 338 (or 10.4 percent) are early decision applicants for whom we do not observe outcomes from regular college applications.

Our basic set of estimates does not use the behavior of these 338 students. (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 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 on our predicting admissions and aid outcomes for the colleges in the 338 students' most preferred college lists that did not result in completed regular admissions processes. We have good conditions for forming these predictions because we nearly always observe outcomes for similarly qualified students *from the same school.* This is a useful consequence of our sample design. We use the following procedure to form the predictions for the "incomplete" colleges. In order to eliminate colleges to which the student would probably not have been admitted, we first eliminate incomplete colleges where the student's own SAT score would put him below the colleges' 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 incomplete 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 package for each student at each incomplete college by using the aid package of another student in his school who *was* admitted to that college. Because grants depend mainly on merit, each incomplete 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 incomplete 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, 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 because they are

based on explanatory variables that are measured with error for the incomplete colleges (attenuation bias).

6.4.5   The Specification of $\beta' \mathbf{x}_{ij}$

Except for reasons of parsimony, we have not attempted to restrict the set of variables that affect college choice. We have included all variables available to us that seemed at all likely to affect students' choices.

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 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. 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 restrict most $\mathbf{x}_{ij}$ variables to having a linear effect because this choice facilitates interpretation. We have, however, allowed some variables, such as distance and a student's SAT "match" with a college, to have nonlinear effects.

6.4.6   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 α in the following expression:

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

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

The way to interpret the odds ratio for a certain variable in $\mathbf{x}_{ij}$ 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 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 being equal; an odds ratio less than one means that an increase in the variable *reduces* a student's probability of matriculating, all else being equal.

The Z statistics are easy to interpret. They are akin to 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.

6.5   How Aid Affects College Choice

6.5.1   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 6.3. 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 6.3's overall pattern of signs and statistical significance, students' college choices appear to be very reasonable. Students are more likely to attend a college if, all else being 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 being equal, its tuition is higher, its room and board is higher, its mean SAT score is 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 are close to our expectations. But are they so close to our expectations when we examine the odds ratios in detail?

The left-hand column of table 6.3 shows us that an additional thousand dollars in grants raises a student's probability of matriculating by 11 percent

	Estimated Odds Ratio
Table 6.3	**The Determinants of College Choice: Estimated Odds Ratios from Conditional Logic Regressions in Which the Binary Outcome Is Matriculation**

	Estimated Odds Ratio
Grant (in thousands), specific to the college	1.108*
	(14.81)*
Loan (in thousands) from the college	1.068*
	(4.03)*
Work study amount (in thousands) from the college	1.125*
	(1.64)*
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	0.980*
	(–1.90)*
College's room and board (in thousands)	0.903*
	(–3.28)*
College's per-pupil instructional spending (in thousands)	1.020*
	(6.02)*
Student's SAT score is this number of percentiles *above* college's average SAT score	0.959*
	(–6.45)*
Student's SAT score is this number of percentiles *below* college's average SAT score	1.001
	(0.011)
College is *most* selective to which student was admitted	1.631*
	(7.41)*
College is *least* selective to which student was admitted	0.694*
	(–4.23)*
Father is alumnus of this college	1.703*
	(3.62)*
Mother is alumna of this college	1.001
	(0.189)
Sibling attended or attends this college	1.896*
	(5.04)*
Distance between college and student's high school, in hundreds of miles	1.000
	(0.06)
Square of distance between college and student's high school, in 10,000s of miles	1.000
	(1.03)
Cube of distance between college and student's high school, in 1,000,000s of miles	1.000
	(–1.01)
College is in-state for the student	1.162
	(1.59)
College is public	1.201
	(1.59)
No. of observations	9,112
Likelihood ratio (chi²)	1,171.41
Prob > chi²	0
Log-likelihood	–2,335.57
Pseudo R^2	0.201

Source: College Admissions Project.

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.

*Statistically significantly different from zero at the 10 percent level.

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 10 percent of his prior probability. These results hint at students' being more sensitive to grants and room and board than to reduced tuition. This is not altogether surprising. A student's reactions to a dollar of grants and tuition need not be *identical.* An increase in his grant affects the individual student's costs but leaves the college environment pretty much unchanged. In contrast, a reduction in tuition lowers the college's revenues, which may reduce the quality of the college environment.

What is more surprising is the response to loans and work study. Recall that an additional thousand dollars in grants raises a student's probability of matriculating by 11 percent. In comparison, an additional thousand dollars in loans raises a student's probability of matriculating by 7 percent of his prior probability, and an additional thousand dollars in work study raises a student's probability of matriculating by 13 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 forgone leisure.

For every percentile that a student's SAT score *exceeds* the mean SAT score of a college, his probability of matriculating falls by 5 percent of his prior probability. Although the odds ratio on the percentile that a student's SAT score *falls 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 63 percent of the prior probability if the college is the most selective among the colleges to which he was admitted, and it falls by 31 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 for students to observe and econometricians to measure than resources are.

The overall resource measure in table 6.3 is the college's per pupil spending on student-related activities (instruction, academic support, student services, and scholarships). Each additional thousand dollars in spending

raises a student's probability of matriculating by 2 percent of his prior probability. This may not sound like much, but it means that a 50 percent increase in spending would raise a student's probability of matriculating by about 31 percent. (The last estimate is for the average college in the sample.)

Finally, having a father or sibling who attended the college greatly increases a student's own probability of attending it. Having a father who attended raises the probability of matriculating by 70 percent of the prior probability; having a sibling who attended raises the probability by 90 percent. 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).

6.6   Are Early Decision Students Different?

In table 6.4, we compare our basic results to results that include early applicants accepted by only one college. Specifically, the left-hand column of table 6.4 repeats our basic results from the left-hand column of table 6.3, 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 6.4.

The table shows that the results change very little when the early decision students are included with their predicted choice sets. The coefficients that change in an interesting way are those of the college's being the least or most selective college in the choice set. Consider the coefficient on the most selective college: The change in it suggests that early decision applicants are less apt than other students to matriculate at the most selective college to which they were admitted. The behavioral interpretation of this result is that early decision applicants are somewhat risk averse and also act strategically: They apply early decision knowing that they will face slightly less stringent admissions criteria. What they give up is the chance to apply to and matriculate at the college that would have been their "long shot."[17]

Now consider the coefficient on the least selective college. Its change suggests that early decision applicants are also less apt than other students to matriculate at the least selective college to which they were admitted. Again, this fits with postulated strategic behavior among early decision applicants: They aim for greater certainty at a college with moderately high selectivity on their list, not for certainty at the least selective college on their list.

17. Another interpretation of the change in the "most selective" coefficient is that we have mistakenly predicted admission to a college that the student himself knew would not admit him. It is unclear how the student would know such a thing if his admissions test scores were, indeed, above the median.

Table 6.4   **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	1.108*	1.102*
	(14.81)*	(14.75)*
Loan (in thousands) from the college	1.068*	1.073*
	(4.03)*	(4.29)*
Work study amount (in thousands) from the college	1.125*	1.113
	(1.64)*	(1.47)
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	0.980*	0.984*
	(–1.90)*	(–1.78)*
College's room and board (in thousands)	0.903*	0.893*
	(–3.28)*	(–3.76)*
College's per-pupil instructional spending (in thousands)	1.020*	1.020*
	(6.02)*	(5.35)*
Student's SAT score is this number of percentiles *above* college's average SAT score	0.959*	0.950*
	(–6.45)*	(–8.45)*
Student's SAT score is this number of percentiles *below* college's average SAT score	1.001	1.003
	(0.011)	(0.26)
College is *most* selective to which student was admitted	1.631*	1.313*
	(7.41)*	(6.57)*
College is *least* selective to which student was admitted	0.694*	0.886*
	(–4.23)*	(–3.49)*
Father is alumnus of this college	1.703*	1.650*
	(3.62)*	(3.54)*
Mother is alumna of this college	1.001	0.966
	(0.18)	(–0.19)
Sibling attended or attends this college	1.896*	1.854*
	(5.04)*	(5.03)*
Distance between college and student's high school, in hundreds of miles	1.000	1.000
	(0.06)	(–0.05)
Square of distance between college and student's high school, in 10,000s of miles	1.000	1.000
	(1.03)	(0.72)
Cube of distance between college and student's high school, in 1,000,000s of miles	1.000	1.000
	(–1.01)	(–0.70)
College is in-state for the student	1.162	1.196*
	(1.59)	(1.95)*
College is public	1.201	1.119
	(1.26)	(0.80)
No. of observations	9,112	10,227
Likelihood ratio (chi²)	1,171.41	1,447.94
Prob > chi²	0	0
Log-likelihood	–2,335.57	–2,516.18
Pseudo R^2	0.201	0.223

Notes: This table is the same as table 6.3, except that the right-hand column substitutes predicted choice sets for actual choice sets for early decision applicants. The prediction procedure is described in the text. All other notes from table 6.3 apply.

*Statistically significantly different from zero at the 10 percent level.

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 onward.[18]

6.7   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 backgrounds are much less similar along the dimensions of family income, parents' college experience, and so on. In tables 6.5 through 6.8, we show the results of reestimating our basic conditional logit specification (the specification in table 6.3) for different subsets of students.

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

For table 6.5, 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 90 percent confidence.

Students from families with low incomes respond to $1,000 in grants by raising their probability of matriculation by about 11 percent of their prior probability. Students with medium-low and medium-high family income respond, respectively, by raising the probability of matriculation by about 13 percent. Students from high-income families also respond but only by raising the probability of matriculation by 8 percent of their prior probability. We can reject the hypothesis that the effect is the same for all income groups with 90 percent confidence. Despite the statistically significant differences, we were somewhat surprised by the *similarity,* not the difference, in the response of meritorious students from different ends of the income spectrum.

We cannot reject the hypothesis that all income groups respond identi-

---

18. Readers may be interested to know that if we use predicted choice sets for *all* students (not just early applicants), we obtain estimates that suffer from attenuation bias. This is what we expect because the aid variables are only estimates for many observations.

Table 6.5 **Are Students from Low- and High-Income Families Equally Sensitive to the Determinants of College Choice? Estimated Odds Ratios from Conditional Logit Regressions**

	Parents' Income				Same effect for all groups?
	Low	Medium Low	Medium High	High	
Grant (in thousands), specific to the college	1.114* (5.15)*	1.128* (7.98)*	1.133* (10.68)*	1.075* (5.02)*	Rejected
Loan (in thousands) from the college	1.036 (0.65)	1.067* (1.68)*	1.072* (2.55)*	1.076* (2.60)*	
Work study amount (in thousands) from the college	1.059 (1.32)	1.219 (1.45)	1.081 (0.66)	1.040 (1.66)	
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	1.000 (0.47)	0.981 (−0.40)	0.964* (−2.27)*	0.980 (−0.57)	
College's room and board (in thousands)	0.840* (−1.67)*	0.945 (−0.85)	0.911* (−1.82)*	0.898* (−1.86)*	
College's per-pupil instructional spending (in thousands)	1.031* (2.56)*	1.016* (1.98)*	1.020* (3.69)*	1.021* (3.89)*	
Student's SAT score is this number of percentiles *above* college's average SAT score	0.986 (−0.82)	0.946* (−3.86)*	0.956* (−4.33)*	0.957* (−3.36)*	
Student's SAT score is this number of percentiles *below* college's average SAT score	0.986 (−0.59)	1.019 (0.77)	1.008 (0.43)	1.015 (0.53)	
College is *most* selective to which student was admitted	1.430 (1.57)	1.612* (3.27)*	1.437* (3.30)*	1.899* (5.41)*	
College is *least* selective to which student was admitted	0.986 (−0.05)	0.798 (−1.18)	0.625* (−3.39)*	0.666* (−2.46)*	

Father is alumnus of this college	3.772	1.293	2.395*	1.256	
	(1.73)	(0.69)	(3.60)*	(0.94)	
Mother is alumna of this college	0.268	2.231*	0.510*	1.919*	Rejected
	(−1.02)	(1.79)*	(−2.13)*	(2.14)*	
Sibling attended or attends this college	2.559*	1.900*	2.268*	1.328	
	(2.04)*	(2.18)*	(4.17)*	(1.18)	
Distance between college and student's high school, in hundreds of miles	0.971	1.006	1.003	0.998	
	(−1.20)	(0.36)	(0.22)	(−0.13)	
Square of distance between college and student's high school, in 10,000s of miles	1.000*	1.000	1.000	1.000	
	(−2.00)*	(0.47)	(−0.08)	(1.17)	
Cube of distance between college and student's high school, in 1,000,000s of miles	1.003*	1.000	1.000	1.000	
	(2.00)*	(−0.47)	(0.09)	(−1.13)	
College is in-state for the student	0.967	1.079	1.379*	1.001	
	(−0.10)	(0.36)	(2.08)*	(0.01)	
College is public	1.602	1.909	1.128	1.062	
	(1.00)	(2.02)	(0.51)	(0.22)	
No. of observations	838	2,011	3,459	2,731	
Likelihood ratio (chi²)	151.91	348.23	456.62	314.11	
Prob > chi²	0	0	0	0	
Log-likelihood	−194.7	−468.1	−884.73	−715.03	
Pseudo R^2	0.28	0.27	0.21	0.18	

Source: College Admissions Project.

Notes: The table shows results from conditional logit estimation of how a student chooses his or her matriculation college among the colleges to which he or she was admitted. The results are shown as odds ratios, with z statistics in parentheses below the odds ratios. 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); and high (greater than or equal to \$140,000). 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.

*Statistically significantly different from 1 with at least 95 percent confidence.

cally to $1,000 of loans, although the point estimates hint that students from low-income families are less attracted by loans—perhaps because they foresee their families having trouble paying off loans. Also, we cannot reject the hypothesis that $1,000 in work study has the same effect on students from all income groups. Moreover, the pattern of estimates on work study is quite difficult to interpret. Students may respond to work study in a heterogeneous way that shows up as coefficients that bounce around. We suspect that this may be because work study has important attributes, contained in the job itself, that we do not observe. For instance, it may be that work study at one college is an academically valuable research job, while work study at another college is an onerous cleaning job.

The greatest negative response to tuition is among students from medium-high-income families. This is not surprising because medium-high-income families are well off enough to pay list tuition but not so well off that list tuition is small relative to their incomes. Low-income parents appear to be quite insensitive to tuition differences. This is probably because they rely on need-based aid and rarely pay the marginal tuition dollar anyway.

As family income rises, students become less willing to attend a college where their own SAT scores exceed the college's mean score. Conversely, students become more eager to attend a college where their own SAT scores fall below the college's mean score. For instance, the draw of the most selective college in a student's choice set rises from a 43 percent increase in the probability of matriculation for the low-income families to a 90 percent increase in probability for the high-income families. The repulsion of the least selective college in the choice set goes from a 1 percent decrease in the probability of matriculation for low-income families to a 33 percent decrease in probability of matriculation for high-income families.

Compared to other students, students from high-income families respond less to having a father or sibling who attended the college but respond to having a mother 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.

6.7.2   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 6.6, where parents are divided into groups based on the selectivity of their colleges (the maximum of the two parents' colleges selectivity).[19] Because the students themselves are solid applicants for *very* selective colleges, we were

---

19. 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.

Table 6.6   **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**

	Parents' College Selectivity			Same effect for all groups?
	Low	Medium	High	
Grant (in thousands), specific to the college	1.120*	1.100*	1.075*	
	(13.15)*	(6.41)*	(3.19)*	
Loan (in thousands) from the college	1.075*	1.096*	1.036	
	(3.48)*	(1.84)*	(0.93)	
Work study amount (in thousands) from the college	0.995	1.354*	2.534*	Rejected
	(−0.06)	(2.00)*	(2.62)*	
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	0.980*	0.965*	1.029*	
	(−1.60)*	(−1.60)*	(0.90)*	
College's room and board (in thousands)	0.882*	1.021	0.834	Rejected
	(−3.17)*	(0.34)	(−1.89)	
College's per-pupil instructional spending (in thousands)	1.018*	1.023*	1.007	
	(4.68)*	(3.46)*	(0.88)	
Student's SAT score is this number of percentiles *above* college's average SAT score	0.962*	0.946*	0.952*	
	(−5.01)*	(−3.65)*	(−2.05)*	
Student's SAT score is this number of percentiles *below* college's average SAT score	1.022	1.176	1.496*	
	(0.71)	(1.96)	(4.74)*	
College is *most* selective to which student was admitted	1.496*	1.897*	1.635*	
	(4.74)*	(4.73)*	(2.60)*	
College is *least* selective to which student was admitted	0.699*	0.653*	0.722	
	(−3.32)*	(−2.31)*	(−1.26)	
Father is alumnus of this college	1.151	1.925*	1.493	
	(0.44)	(2.82)*	(1.45)	
Mother is alumna of this college	1.342	0.780	0.861	
	(0.84)	(−0.83)	(−0.37)	
Sibling attended or attends this college	1.936*	2.455*	0.841	
	(4.22)*	(3.43)*	(−0.39)	
Distance between college and student's high school, in hundreds of miles	1.010	0.878*	1.038	Rejected
	(1.12)	(−4.57)*	(1.66)	
Square of distance between college and student's high school, in 10,000s of miles	1.000	1.000	1.000	
	(0.53)	(2.18)	(−0.02)	
Cube of distance between college and student's high school, in 1,000,000s of miles	1.000	1.000	1.000	
	(−0.53)	(−1.43)	(−0.03)	
College is in-state for the student	1.191	0.688*	2.110*	Rejected
	(1.46)	(−1.87*)	(2.59)*	
College is public	1.206	1.587	0.789	
	(1.04)	(1.53)	(−0.48)	
No. of observations	5,673	2,280	1,159	
Likelihood ratio (chi²)	733.27	362.49	171.52	
Prob > chi²	0	0	0	
Log-likelihood	−1,447.93	−555.72	−283.98	
Pseudo R^2	0.2	0.25	0.23	

Source: College Admissions Project.

Notes: The table shows results from conditional logit estimation of how a student chooses his or her matriculation college among the colleges to which he or she was admitted. The results are shown as odds ratios, with z statistics in parentheses below the odds ratios. 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); and 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 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.

*Statistically significant different from 1 with at least 95 percent confidence.

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.[20]

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 percent of the prior priority for a student with low-selectivity parents, but the corresponding number is only 8 percent for a student with high-selectivity parents. Also, $1,000 in loans raises the probability of matriculation by 8 percent of the prior priority for a student with low-selectivity parents, but the corresponding number is only 4 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 interpret this result literally because of the variation in work-study jobs.

The difference in the response to tuition is interesting: 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 to 3 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 3 percent of their prior probability). It is unlikely that tuition itself is attractive to the high-selectivity parents, but it is quite probably correlated with measures of college resources that are not in the regression (for instance, the spending dedicated to undergraduates alone or nonlinear effects of per-pupil spending on students).

We also find that, compared to other students, students with high-selectivity parents are less attracted to their sibling's college. Students with medium-selectivity parents appear to be the most attracted by their father's and sibling's college.

6.7.3   Students from Public and Private Secondary Schools

In table 6.7, we investigate whether students from public and private high schools make college choices differently. We find that an extra $1,000

---

20. 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 a 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.

Table 6.7   **Are Students from Private and Public High Schools Equally Sensitive to the Determinants of College Choice? Estimated Odds Ratios from Conditional Logit Regressions**

	Public High chool	Private High School	Same effect for all groups?
Grant (in thousands), specific to the college	1.112*	1.106*	
	(11.34)*	(9.49)*	
Loan (in thousands) from the college	1.084*	1.049*	
	(3.89)*	(1.76)*	
Work study amount (in thousands) from the college	1.187*	1.048	
	(1.84)*	(0.43)	
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	0.977*	0.998	
	(−1.78)*	(−0.06)	
College's room and board (in thousands)	0.812*	0.955	Rejected
	(−4.77)*	(0.32)	
College's per-pupil instructional spending (in thousands)	1.020*	1.018*	
	(4.60)*	(3.99)*	
Student's SAT score is this number of percentiles *above* college's average SAT score	0.964*	0.952*	
	(−4.51)*	(−4.53)*	
Student's SAT score is this number of percentiles *below* college's average SAT score	0.992	1.015	
	(−0.51)	(0.83)	
College is *most* selective to which student was admitted	1.685*	1.583*	
	(5.70)*	(4.74)*	
College is *least* selective to which student was admitted	0.781*	0.555*	Rejected
	(−2.16)*	(−4.31)*	
Father is alumnus of this college	1.971*	1.470*	
	(3.43)*	(1.70)*	
Mother is alumna of this college	0.785	1.406	
	(−0.97)	(1.14)	
Sibling attended or attends this college	2.176*	1.388	Rejected
	(4.95)*	(1.48)	
Distance between college and student's high school, in hundreds of miles	1.004	0.987	
	(0.41)	(−1.10)	
Square of distance between college and student's high school, in 10,000s of miles	1.000	1.000	
	(0.13)	(1.15)	
Cube of distance between college and student's high school, in 1,000,000s of miles	1.000	1.000	
	(−0.13)	(−1.04)	
College is in-state for the student	1.301*	0.989	
	(2.02)*	(−0.07)	
College is public	0.967	1.96*	Rejected
	(−0.18)	(2.88)*	
No. of observations	4,817	4,295	
Likelihood ratio (chi²)	662.95	555.86	
Prob > chi²	0	0	
Log-likelihood	−1,232.24	−1,079.63	
Pseudo R^2	0.21	0.20	

Source: College Admissions Project.

Notes: The table shows results from conditional logit estimation of how a student chooses his or her matriculation college among the colleges to which he or she was admitted. The results are shown as odds ratios, with *z* statistics in parentheses below the odds ratios. 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.

*Statistically significantly different from 1 with at least 95 percent confidence.

in room and board reduces the probability of matriculating by 19 percent of prior probability among public school students, but that it has a statistically insignificant effect of 5 percent 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 21 percent of the prior probability. For private school students, the corresponding number is a much larger—44 percent. Being an in-state college is an attraction for public school students, whose probability of matriculation rises by 30 percent of their prior probability. Being in-state has no such draw for private school students. Interestingly enough, a college's being public attracts private school students but has no such effect on public school students.

6.7.4 Summing up the Differences in College Choice among Students from Different Backgrounds

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 several other dimensions: gender; region of the country; recipiency of an outside scholarship, like the National Merit Scholarship; a record of leadership while in high school; size of the high school; and so on. There are no statistically significant differences by race either, but this may because there are insufficiently few nonwhites in the sample to extract distinct patterns from their choice behavior. We suspect that high aptitude 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).

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 6.5 through 6.7 really add up to much. One way 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 6.8, we use the low-income students' coefficients with the high-income students' data and vice versa. We performed the 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 pre-

Table 6.8 **Share of Students Who Would Matriculate at Another College if They Were to Obey the College Choice Model of Students from a Different Background**

	Parent Income		Parents' College Selectivity		High School	
	Actually Low[a]	Actually High[b]	Actually Low[c]	Actually High[d]	Actually Public[e]	Actually Private[f]
Share who would matriculate at a different college in their choice set	0.315	0.417	0.244	0.238	0.167	0.172

Source: Predictions based on conditional logit regressions shown in tables 6.4 through 6.6.
[a]For students acting like high-income students.
[b]For students acting like low-income students.
[c]For students acting as though it had been high.
[d]For students acting as though it had been low.
[e]For students acting like private school students.
[f]For students acting like public school students.

dicted 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, 42 percent of whom would matriculate at a different college if they made choices the way that low-income students do. Similarly, 32 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 17 percent when we "exchange" the models of public and private high school students and are in the range of 24 percent when we exchange the models of students whose parents attended low- and high-selectivity colleges.

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 this chapter. The reason we cannot give a satisfactory answer is that we are constraining students to rechoose 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 6.4 through 6.6, this constraint is 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 manner, however, because they are relative ones: the share of students 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, 95 percent of high-income students choose the most selective college in their choice set when they act like themselves, but only 57 percent would do so if they were to act like low-income students. Eighteen percent of low-income students choose the least selective college in their choice set when they act like themselves, but only 7 percent would do so if they were to act like high-income students.

6.8   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 similarly to every additional $1,000, regardless of whether it was a grant or a loan. 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 responses to grants, their responses to loans and work study are too large.

Furthermore, tables 6.5 through 6.7 show that, while some students (high-income, high selectivity of parents' college, 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 and loans about equally.

6.8.1   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 largely ignore such front-loading and 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 spending on students. Obviously, if all colleges had similar tuition, comprehensive cost, and 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 per-pupil spending on students is $15,489. Students in the sample applied to forty colleges at which per-pupil spending on students was less than $7,000, and they applied to about an equal number of colleges at which per-pupil spending on students was more than $28,000, which is *four times* $7,000. 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, and the correlation between the amount of a grant and the percentage of comprehensive costs that it represents is 0.467.

If students behave according to a standard model of human capital in-

21. We do not assume that students 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 PhD programs. The same considerations do not apply to undergraduate programs.

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, nonflagship 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).

vestment, 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 a measure of the investment made freely *by others* in a student's human capital.[23] Given the amount of a grant, a student will care about the grant's *share* of tuition or costs only if he is an irrational investor (for instance, naively flattered by receiving a large *share* of tuition) or 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 $28,000 expenditure truly represents a smaller investment than 100 percent of a $7,000 expenditure.

6.8.2   Evidence on How Students Respond to the Aid Variables They Should Ignore

Table 6.9 presents our estimates of how students make college choices when we allow them to respond to aid variables they should ignore. Column (1) of the table reproduces the estimates in table 6.3. Recall that the estimates suggest that each additional $1,000 of a grant is estimated to increase the probability of matriculation by 11 percent of the prior probability.

In column (2), 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 (2) 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 (2) also shows that, for a grant of a given amount, calling it a scholarship increases the probability of matriculation by 86 percent of the prior probability. This is a great effect for 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 a nonlinear effect of the 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 48 percent of the prior probability. Again,

---

23. There is another measure of the investment made *by others* in a student's human capital. It is the implicit grant created by the difference between a college's per-pupil expenditure and its list tuition. In practice, the size of a college's implicit grant increases with its selectivity.

Table 6.9   **Do Aid Variables That Should Not Matter Affect College Choice? Estimated Odds Ratios from Conditional Logit Regressions**

	(1)	(2)	(3)
Grant (in thousands), specific to the college	1.108*	0.968	0.968
	(14.81)*	(–1.46)	(–1.45)
Loan (in thousands) from the college	1.068*	1.062*	1.060*
	(4.03)*	(3.67)*	(3.54)*
Work study amount (in thousands) from the college	1.125*	1.207*	1.204*
	(1.64)*	(2.53)*	(2.48)*
College's tuition (in thousands), in-state or out-of-state as appropriate to the student	0.980*	0.998	0.998
	(–1.90)*	(–0.65)	(0.50)
College's room and board (in thousands)	0.903*	0.943*	0.946
	(–3.28)*	(–1.70)*	(–1.62)
College's per-pupil instructional spending (in thousands)	1.020*	1.020*	1.018*
	(6.02)*	(6.05)*	(5.64)*
Grant is this share of college's per-pupil instructional spending			1.121
			(0.50)
Grant is called a "scholarship"		1.860*	1.838*
		(6.15)*	(6.00)*
Grant is front-loaded (more in freshman year)		1.479*	1.475*
		(1.90)*	(1.88)*
Grant is this share of college's tuition		0.792	0.801
		(–0.98)	(–0.91)
Grant is this share of college's comprehensive cost		27.551*	23.511*
		(3.70)*	(3.30)*
Student's SAT score is this number of percentiles *above* college's average SAT score	0.959*	0.952*	0.951*
	(–6.45)*	(–7.31)*	(–7.36)*
Student's SAT score is this number of percentiles *below* college's average SAT score	1.001	0.998	0.997
	(0.11)	(–0.27)	(–0.18)
College is *most* selective to which student was admitted	1.631*	1.641*	1.644*
	(7.41)*	(7.38)*	(7.39)*
College is *least* selective to which student was admitted	0.694*	0.676*	0.664*
	(–4.23)*	(–4.43)*	(–4.60)*
Father is alumnus of this college	1.703*	1.629*	1.651*
	(3.62)*	(3.21)*	(3.26)*
Mother is alumna of this college	1.001	1.013	0.988
	(0.19)	(0.06)	(–0.06)
Sibling attended or attends this college	1.896*	1.907*	1.975*
	(5.04)*	(4.99)*	(5.21)*
Distance between college and student's high school, in hundreds of miles	1.000	1.003	1.003
	(0.06)	(0.36)	(0.42)
Square of distance between college and student's high school, in 10,000s of miles	1.000	1.000	1.000
	(1.03)	(1.20)	(1.20)
Cube of distance between college and student's high school, in 1,000,000s of miles	1.000	1.000	1.000
	(–1.01)	(–1.19)	(–1.19)
College is in-state for the student	1.162	1.191*	1.220*
	(1.59)	(1.80)*	(2.03)*
College is public	1.201	1.513*	1.536*
	(1.26)	(2.73)*	(2.80)*
No. of observations	9,112	9,112	9,112
Likelihood ratio (chi²)	1,171.41	1,283.68	1,275.08
Prob > chi²	0	0	0
Log-likelihood	–2,335.57	–2,255.89	–2,232.17
Pseudo R^2	0.2	0.22	0.22

Source: College Admissions Project.

Notes: The table shows results from conditional logit estimation of how a student chooses his or her matriculation college among the colleges to which he or she was admitted. The results are shown as odds ratios, with *z* statistics in parentheses below the odds ratios.

*Statistically significantly different from zero with at least 95 percent confidence.

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 previously ($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 $4,435 per year. While we should not take these numbers literally (because they are require a good deal of extrapolation and are based on different columns of table 6.9), it is fairly clear that students respond excessively to front-loading as compared to the amount of the grant.

The next two rows of column (2) 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 275.5 percent of his prior probability. At first glance, the odds ratio may be implausibly large, but let us interpret it. Suppose that a student's prior probability of matriculation was 10 percent. Then, his posterior probability would be 27.5 percent with a grant of 10 percent of comprehensive costs and 55.1 percent with a grant of 20 percent of comprehensive costs. What the high odds ratio is telling us 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.[24] Column (3) simply repeats the specification of column (2), adding the share of per-pupil spending that the grant represents. This additional variable does not have a statistically significant coefficient, but the magnitude of the odds ratio on the grant's share of comprehensive costs does fall somewhat (by about 15 percent).

Why is it that the grant's share of *comprehensive cost* matters, rather than its share of tuition or per-pupil spending on students? 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 un-

---

24. In addition, we should not really focus exclusively on the odds ratio attached to 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.

able to pay cash). The grant's share of spending is only weakly related to either motive. Spending on students is difficult for students to observe, so the grant's share of spending is unlikely to be sufficiently salient to be flattering. Also, the grant's share of 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 both more salient and more relevant to the cash constrained. After all, there is a 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; $2,000; $2,500; etc.). At these colleges, the nontuition part of comprehensive cost is three to four times as large as tuition and makes up the lion's share of out-of-pocket expenses.

6.8.3   Do Students' Responses to Aid Variables That Should Not Matter Depend on Their Backgrounds?

In tables 6A.3 through 6A.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 reestimate the regression shown in table 6.9 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 6.9, tables 6A.3 through 6A.5 present only the 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 whose 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 medium-low-income students. 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 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 substantial as the payment asked of medium-high and high-income students). Yet medium-low income parents may be unwilling 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 represent a large share of comprehensive costs.

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

6.9   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 behaving like rational investors in their own human capital. That is, are students making the best use of aid in order to maximize their lifetime economic well-being? Thus far, we have found two fairly obvious violations of 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. Remember our example in which a student could be admitted to colleges A, B, and C, and where college C was 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. If college C offered more aid, then no trade-off would exist: The student would do better all around by matriculating at college C. We argued that this no-trade-off situation cannot hold generally because it would not be an equilibrium: College C would be so oversubscribed that it would automatically become more selective so that the student admitted to A and B might no longer be admitted to C.

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 generally exist. 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 spending on students 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.

6.9.1   An 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: (1) 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 reduced human capital; and (2) 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 decrease in future earnings, which will be caused by reduced human capital.[25]

We need to measure the extent to which students reduce their human capital investment in return for increased aid. To make such a measurement, we use the estimated college choice model from the previous sections and 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' predicted college choices change. That is, we see how students' choices would predictably change in the absence of aid.[26]

Once we have measures of how students' predicted college choices change when we remove aid, we can estimate the losses associated with the

25. 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.

26. Note that we look at how students' *predicted* college choices change when we zero out aid. Thus, we do not incorrectly attribute to aid the differences between students' actual and predicted behavior.

reductions in human capital and consumption that they accept in return for aid. We can then compare these losses to the value of the aid we removed.

6.9.2   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: we could measure the *inputs* available at each college or we could measure the *value added* of each college. If we wanted to measure the value added, we would need to compute the earnings associated with each college and then control for differences in earnings due to differences in students' incoming aptitudes. We would not want to attribute all of a student's earnings to his college; much of his earnings would be due to the abilities that got him admitted to the college in the first place.

Although—in theory—either the inputs or value added strategy could be pursued, we reject the value added strategy as impractical. There have been several attempts to measure the value added associated with colleges, but there are no commonly accepted estimates and some of the best-known estimates are deeply flawed.[27] However, we do not reject the value added strategy to avoid controversy. We reject the strategy because every researcher would agree that it is currently impossible to estimate value added for a *wide range* of *specific* colleges. Some might argue that we could estimate value added for a handful of specific colleges; some might argue that we could estimate value added for coarse groups of colleges (groups so coarse that some would include hundreds of colleges). No one would argue that we could estimate value added for many specific colleges. This is for a

---

27. For instance, Dale and Krueger (1999) attempted to estimate the return to attending specific colleges in the College and Beyond survey data. They assigned individual students to a "cell" based on the colleges to which they are admitted. *Within* a cell, they compared those who attend a more-selective college (the treatment group) to those who attended a less-selective college (the control group). If this procedure had gone as planned, all students within a cell would have had the same menu of colleges and would have been arguably equal in aptitude. The procedure did not work in practice because the number of students who reported more than one college in their menu was very small. Moreover, among the students who reported more than one college, there was a very strong tendency to report the college they attended plus one less-selective college. Thus, there was almost no variation within cells if the cells were based on actual colleges. Dale and Krueger (1999) were forced to merge colleges into crude "group colleges" to form the cells. However, the crude cells made it implausible that all students within a cell were equal in aptitude, and this implausibility eliminated the usefulness of their procedure. Because the procedure works best when students have large menus, and most students do not have such menus, the procedure essentially throws away much of the data. A procedure is not good if it throws away much of the data and still does not deliver "treatment" and "control" groups that are plausibly equal in aptitude. Put another way, it is not useful to discard good variation in data without a more than commensurate reduction in the problematic variation in the data. In the end, Dale and Krueger (1999) predictably generate statistically insignificant results, which have been unfortunately misinterpreted by commentators who do not have sufficient econometric knowledge to understand the study's methods.

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.[28] 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 are choosing within a single coarse group.

Therefore, we use the inputs strategy for measuring human capital investment. Our approach to measuring inputs is conservative by design: We count instructional spending and *only* instructional spending as human capital investment. Instructional spending not only excludes spending on research, it also excludes some categories of spending on students: student services (such as health care), academic support, and scholarships. In other words, a good deal of spending at high-spending colleges is excluded, even though we believe that much of the excluded spending is complementary to instructional spending and produces greater human capital.[29] By focusing on instructional spending, we "bend over backward" in favor of finding that human capital investment is almost as great at low-spending colleges as it is at high-spending colleges. Empirically, instructional spending is a *much* larger share of total spending at low-spending colleges than it is at high-spending colleges. Among the colleges in our sample, the colleges with the highest per-pupil spending have instructional spending shares around 0.1; the colleges with the lowest per-pupil spending have instructional spending shares around 0.6.

It is evident from the students' own choice behavior that they prefer more-able peers, and it seems likely that peers should be regarded as inputs. That is, part of the human capital gained by a student is probably generated by peer spillovers or by the interaction of good peers and college re-

28. 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 U.S. 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, 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.

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

sources. There is no simple way to quantify peer inputs and add them to instructional inputs. Nevertheless, we do not wish to ignore peer inputs and recognize only instructional inputs. We deal with this problem in a way that is at least transparent: We simply show changes in peers' SAT scores and remind readers that they should mentally add peer inputs with the weight they consider appropriate.

Although the students' choice behavior suggests 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, others (such as faculty time) can be disproportionately allocated to certain students. Logically, the disproportionality must be a function of the degree to which a high-aptitude student 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 for themselves whether more able peers are not generators or net destroyers of human capital, for a given level of instructional spending.

6.9.3   How Students' College Choices Would Differ in the Absence of Aid, Part 1: The Structure of Table 6.10

To create table 6.10, we first estimate the conditional logit model shown in the second column of table 6.9. We then use the coefficient estimates to predict which colleges the student would be most likely to attend (1) with the aid he was actually offered; and (2) in the absence of aid (that is, with all of the aid variables zeroed out).[30] For simplicity, we will call the former college the "with-aid college" and the latter college the "without-aid college." We then compute the present value of the aid we zeroed out, and we show this in column (2) of table 6.10.[31] That is, column (2) shows the benefit of taking aid. The succeeding columns show the costs of taking aid. Column (3) shows the difference in consumption between the with-aid and without-aid colleges. We measure the difference in consumption by sub-

30. We allow the student to rechoose among all of the colleges that appear in the College Admissions Project sample. 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 3,000 other institutions of higher education in the United States that virtually never enroll a student like those in our sample. They include community colleges and other institutions that never or rarely grant the baccalaureate degree.

31. 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.

Table 6.10   Students Who Gain and Lose Lifetime Present Value by Responding to Aid

	% of Students in Group (1)	Average Value of Aid ($) (2)	Average Change in College Consumption, Due to Student's Response to Aid ($) (3)	Average Change in Instructional Spending, Due to Student's Response to Aid ($) (4)	Average Change in Median Peer's SAT Percentile, Due to Student's Response to Aid (5)	Average Change in Lifetime Present Value, Due to Student's Response to Aid ($) (6)
Students who gain lifetime present value, due to their response to aid	30.8	11,534	−171	−26	−2.7	44,075
Students who lose lifetime present value, due to their response to aid	38.9	2,793	−960	−14,538	−8.5	−76,096

Source: Based on predictions from conditional logit equations like that estimated in column (2) of table 6.9. See text for details.

Notes: Top panel students may re-choose among all colleges when all aid is zeroed out. The remaining 30.3 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.

tracting the room and board at the without-aid college from the room and board at the with-aid college. 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.[32]

Column (4) shows the difference in instructional spending between the with-aid and the without-aid colleges. Column (5) shows the difference in the median SAT score, in percentiles, between the with-aid and the without-aid colleges. This difference indicates the change in a student's peer group.

In order to compute the change in a student's net present value from taking aid, we have to make a few assumptions to create the present values. We assume a *real* discount rate of 3 percent per year and an annual inflation rate of 3 percent. We assume a conservative 7 percent real rate of return on human capital investment, and we assume that human capital pays out for forty years.[33] In column (6), we show the change in a student's net present value, due to his response to aid. To get column (6), we first sum the instructional spending difference between the with-aid and without-aid colleges over four years of college. This gives us the human capital asset that pays out at 7 percent for forty years. We compute the present discounted value of this stream of payments and then add the present value of the change in aid and the present value of the change in consumption. Keep in mind that the calculation omits the benefits and costs of peers, which we cannot quantify accurately.

For display in table 6.10, we divide students into two groups. The top row contains students who, when we conducted our thought experiment, appeared to have made good use of the aid they were offered. Their with-aid choices have higher lifetime values than their without-aid choices. 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 an amount exactly equal to the present value of the aid itself. If high achiev-

---

32. If all grants were tuition discounts, it might be appropriate to consider aid as one side of the trade-off and human capital investment (and only human capital investment) as the other side of the trade-off. But, 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.

33. Our assumption about the number of years over which human capital pays out is not crucial because the out years are so heavily discounted. Our assumption about the rate of return to human capital is more important, so we make a conservative assumption of 7 percent, which is near the bottom of the generally accepted range of estimates. It is probably especially conservative for the highly meritorious group of students whom we are studying. Indeed, the tendency of highly meritorious students to continue in school beyond the baccalaureate degree strongly suggests that they earn a supernormal rate of return during their baccalaureate years, which prompts them to continue enrolling until their rate of return is more in line with their discount rate.

ing peers make a student get more human capital from instructional spending, this group's size is overstated: Some students who appear to have made a rational human capital investment actually gave up too much in the way of peers. If high-achieving peers make a student get less human capital from instructional spending, this group's size is understated.

The bottom row contains students who, when we conducted our thought experiment, appeared to have been seduced by aid into making "irrational" human capital investments. That is, in return for aid, they accepted such large reductions in human capital investment and consumption that they lost lifetime value. If high-achieving peers make a student get more human capital from instructional spending, this group's size is understated. If high-achieving peers make a student get less human capital from instructional spending, this group's size is overstated.

Of course, there is a third group of students: 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. 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 be obtained by any student in our sample, no matter how rich he is, if he is willing to attend a less-selective college.

6.9.4   How Students' College Choices Would Differ in the Absence of Aid, Part 2: The Evidence in Table 6.10

Table 6.10 shows that 30.8 percent of students responded to aid in such a way that their lifetime value was increased. Some of these students simply accepted aid at the same college that they would have picked if no aid had been offered. More interestingly, some of these students accepted aid 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: $11,534 per year. They attended a with-aid college that offered consumption that was, on average, $171 lower per year and instructional spending that was, on average, $26 lower per year. These losses are small. As a consequence, by responding like a rational investor to aid, the average student in this group gained lifetime present value of $44,075. This gain would be somewhat different if peers matter because we have not deducted or added any amount for the students' worse peers. The students probably lost some human capital investment because their peers were slightly worse (2.7 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 instructional spending.

The bottom row of table 6.10 shows that 38.9 percent of all students did

not act like rational investors. These students 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: $2,793 per year. Compared to the without-aid college they would have attended, students in this group attended a with-aid college that offered consumption that was, on average, $960 lower per year and instructional spending that was, on average, $14,538 lower per year. On net, the average student in this group lost lifetime present value of $76,096. The loss would be somewhat different if peers matter because we have not deducted or added any amount for the student's worse peers. They probably also lost some human capital investment because their peers performed 8.5 percentile points worse on the SAT.

Readers may be initially surprised that such a nonnegligible share of students lose when they respond to aid, but the statistics in table 6.10 are really a straightforward implication of the behavior that we saw illustrated in table 6.9. The students who lose the most are precisely those students who accept aid that is actually quite modest in value but covers a large share of comprehensive cost at a college that spends very little on instruction. We know from tables 6A.3 and 6A.4 that not all students are equally likely to be 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 cost 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 than other students or because they are less credit constrained than other students.

6.10   Interpreting the Evidence

Overall, we would describe the college choice behavior of the high-aptitude students in our sample as 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, to whether it is in-state, and to whether it is public. However, they are sensitive to tuition, room, and board in the expected direction (lower is better). They also 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 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 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 61.1 percent of students in our sample respond as rational investors would in the presence of aid. Of these rational types, about half do not get enough aid to measurably respond to it, and about half improve their lifetime present value by accepting an aid offer that is more than generous enough to offset the reductions in college consumption and instructional spending associated with the aid. However, about 38.9 percent of students in our sample respond to aid in such a way that they reduce their own lifetime present value. They accept an aid offer that is too small to offset the reductions in consumption and instructional spending that they experience. 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 as a rule. They all have records that show that they can work hard now in return for gains in the distant future. Credit constraints are also not a good explanation for the attractiveness of front-loading, because 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 what 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, is the problem. Over and over, these parents complain that they are baffled by the aid process. They argue that the colleges do not explain their offers well. They complain that other families are more "in the

know." Most of all, they worry about whether their children will benefit sufficiently from greater resources to justify the additional cost. Credit constraints do not receive nearly as much comment: Among parents who commented and whose children exhibited self-defeating responses to aid, only 6.9 percent stated that they were simply unable to pay the costs associated with their child's most preferred college. We do not want to overinterpret the anecdotal evidence from parents' comments because they may have been embarrassed to say that family circumstances prevented 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.

We began this chapter by asking whether highly meritorious students, who are the big investors in the human capital market, act in a manner consistent with maximizing their returns—and thus, American economic growth, which is increasingly dependent on human capital investments. We come down with a very qualified yes: High-aptitude students understand the incentives that they face well enough to "get the sign right" when they react to any one factor. However, a substantial minority of them make trade-offs among factors that are wrong. About a third of the students are probably underinvesting and our conservative calculations suggest that a typical mistake is worth \$76,096 in present value. Being the cost of mere error, \$76,096 is a useful number to keep in mind when thinking about the magnitude of human capital investments and the consequent importance of getting them right.

Appendix

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

(A1) $$\frac{\text{Prob}(\text{collegechoice}_i = j)}{\text{Prob}(\text{collegechoice}_i = j')}$$

is independent of the probabilities of the remaining alternatives in the choice set. This property, the independence of irrelevant alternatives (IIA), is violated in certain applications. For our application, it is probably most useful to think about violations being likely to occur if is a natural nesting structure in students' choice sets. In this section, we first construct an example of a violation in order to elucidate the problem. We then construct an example in which IIA is unlikely to be violated; this example will help readers see how students' endogenous formation of their choice sets is actually helpful. We finally construct an example in which IIA is likely to be violated.

There is a specification test for IIA, but we have decided to explain the

issue logically rather than merely present the results of the test because we believe that a logical understanding will better enable readers to judge our results. Nevertheless, knowing the results of the specification test is helpful. Hausman and McFadden (1984) and McFadden (1987) propose 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 lost. Our application never comes close to being rejected by this specification test: If we remove a random college from the choice sets, we get a test statistic that averages 0.43 (p-value 0.48, it is distributed at χ^2_1). This test statistic is based on the specification in table 6.3.

Consider the following example in which IIA is violated. Suppose that Colgate University (a private liberal arts college) and Ohio State University (a public research university) are in a student's choice set and that we are considering the ratio of the probabilities associated with these institutions. Suppose also that the grant at the University of Michigan (another public research university) increases, raising the probability associated with it. The Colgate–Ohio State probability *ratio* need not change: Michigan becomes a more probable choice overall, both Colgate and Ohio State necessarily become less probable choices overall, but the *relative* probability of Colgate to Ohio State may remain unchanged. Let us say, however, that the student actually had a nested structure to his choice: He first chose his favorite public research university and his favorite private liberal arts college, and then he held a runoff between the top schools from each group. Say that the grant increase makes Michigan bump Ohio State from its place as the student's favorite public research university. Then, the grant not only raises the probability of Michigan overall, it dramatically changes the Colgate-Ohio State probability ratio. IIA is violated; the role that nesting plays becomes clear.

Now consider how students' endogenous selection of their choice sets is helpful. Suppose students use all of the information on colleges that can be observed or predicted at the time of application. Suppose, moreover, that students assume that *unpredictable* college attributes (for instance, the part of an aid package that cannot be predicted based on a college's policies and conventions) are pretty much the same within each of several nesting groups. For instance, a student might figure that if his application has special appeal for private liberal arts colleges, each of them will offer him an aid package that is 5 percent more generous than what they would otherwise offer someone with his characteristics. The student's suppositions produce a natural nesting structure. The student should examine all colleges, apply to the top institution in each nesting group, and wait for the unpredictable attributes to be resolved—as they will be, when he receives his admissions offers. The student can then conduct his runoff.

If the above assumptions and behavior are fulfilled, then the student is

likely to satisfy IIA because his runoff menu of colleges will not contain any nesting structure. He has already taken the nesting structure into account when deciding where to apply; he has already eliminated less preferred alternatives from each nesting group. This leaves us with an endogenous choice set in which it is more plausible that IIA holds. In our example, Michigan's and Ohio State's attributes have already been fully considered and only one of the two institutions is still in the choice set. In the real world, students may not *strictly* obey the above assumptions and behavior, but students do act in accordance with them to a great extent. For instance, no student applies to all or even many public research universities: He applies his nesting structure and eliminates many of the choices within each nest group. We econometricians need not know what the nesting structure is. The student has already applied it; we econometricians need only observe how he makes choices in the runoff among nesting groups.

Let us consider how the previous example may fall afoul of IIA. We assumed that *unpredictable* college attributes are pretty much the same within each of several nesting groups. This assumption is most likely to be violated if, say, the student believes that admission is random within a nesting group. Such randomness is most likely to occur at the top handful of selective colleges, where admissions probabilities are so low that there is probably some arbitrariness in admissions even among institutions that have the same preferences about students. Thus, a student might apply to the top institution in each of his nesting groups, *except,* say, for the nesting group that contains the most selective private research universities. He might, in this one nesting group, figure that he has a 33 percent chance of getting into each of his five favorite institutions and figure that these chances are independent. The student might decide to apply to all five favorites: This would give him an 86.5 percent chance of getting into at least one institution, a 32.8 percent chance of getting into exactly two, and so on. If he gets into multiple institutions in the nesting group (as he will with some nonnegligible probability if he applies to all five), the student will face an endogenous choice set that has some nesting structure.

Overall, we believe that our data do not reject in the IIA specification test because endogenous formation of choice sets works for us. That is, we believe that our data fit the model's restrictions better than data would fit them if we knew nothing about students' endogenous choice sets and estimated a conditional logit with *all* colleges in every choice set. (Indeed, if we put all College Admissions Project colleges into each choice set, we get rejections in the IIA specification test.) We do not claim to have remedied the IIA issue that arises with conditional or multinomial logit estimation, but we believe that students' endogenous choice set formation is largely a help, not a hindrance.

Readers interested in endogenous choice sets where the choice is observed (as it is in our exercise) may wish to consult Peters, Adamowicz, and

Boxall (1995), Haab and Hicks (1997), Hicks and Strand (2000), and Parsons, Plantinga, and Boyle (2000). Manski (1977) considers the case in which endogenous choice sets are not observed so that the econometrician must jointly estimate the endogenous choice set and the choice within the set. Identification is theoretically possible but very difficult to achieve unless there is ancillary evidence with which to predict a person's endogenous choice set.

Finally, 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"). We have been asked whether this implies that students who have more colleges in their choice sets exercise disproportionate influence over the estimates. The answer is "no" on both counts. This becomes clear if we return to the conditional logit equations:

$$\ln L = \sum_{i=1}^{n} \left[\sum_{j=1}^{J_i} \text{matric}_{ij} \ln \text{Prob}(\text{collegechoice}_i = j) \right],$$

$$\text{Prob}(\text{collegechoice}_i = j) = \frac{e^{\beta' \mathbf{x}_{ij}}}{\sum_{j=1}^{J_i} e^{\beta' \mathbf{x}_{ij}}}.$$

Examine the log likelihood equation. It shows that each student contributes equally to the log likelihood because each student's college choice probabilities must sum to 1 (see second equation).

Table 6A.1   **Description of the Colleges to Which Students Applied, from the College Admission Project Data**

Variable	Mean	Standard Deviation	Minimum	Maximum
Matriculated at this college	0.1813	0.3853	0	1
Admitted to this college	0.6566	0.4748	0	1
Applied early to this college	0.1281	0.3389	0	2
Withdrew application from this college, usually after early decision elsewhere	0.0516	0.2212	0	1
Grants specific to this college	1,777.8140	4,933.3550	0	36,000
Loans from this college	413.4718	1,855.6370	0	36,548
Work study amount from this college	110.7380	482.5519	0	15,000
Grant is called a "scholarship"	0.1291	0.3354	0	1
Grant is front-loaded (more in freshman year)	0.0137	0.1161	0	1
Grant is this share of tuition	0.1229	0.3676	0	8
Grant is this share of comprehensive cost	0.0722	0.1902	0	2
Student was a recruited athlete at this college	0.0327	0.1779	0	1
Father is an alumnus of this college	0.0314	0.1744	0	1
Mother is an alumna of this college	0.0209	0.1431	0	1
Sibling attended or attends this college	0.0388	0.1932	0	1
College is public	0.2631	0.4403	0	1
College is private not-for-profit	0.7328	0.4436	0	1
College is international, except for Canadian colleges, which are treated as U.S. colleges	0.0040	0.0633	0	1
College's median SAT score, in national percentiles	83.8816	12.0390	14	98
Student's SAT score is this many percentiles *above* college's median SAT score	8.7393	9.5927	0	82
Student's SAT score is this many percentiles *below* college's median SAT score	1.7454	5.6654	0	68
In-state tuition	18,181.2300	9,198.9780	0	27,472
Out-of-state tuition	20,497.7600	5,890.7530	0	27,472
Tuition that applies to this student	19,276.9000	7,965.1400	0	27,472
Room and board at this college	6,975.7190	1,244.3320	0	10,299
In-state comprehensive cost of this college	25,745.7900	9,935.6770	0	35,125
Out-of-state comprehensive cost of this college	28,059.7200	6,681.4230	0	35,125
Comprehensive cost that applies to this student	26,841.9800	8,662.0230	0	35,125
Per-pupil expenditure on students (instruction, student services, academic support, scholarships) of this college, in thousands	29.9219	17.1009	2	146
Instructional per-pupil expenditure of this college, in thousands	19.8160	12.5401	2	72
College is in-state	0.2666	0.4422	0	1
Distance between student's high school and this college, in miles	673.2152	873.1788	0	5,774
College is in AK	0.001	0.0106	0	1
College is in AL	0.0038	0.0613	0	1
College is in AR	0.0003	0.0168	0	1
College is in AZ	0.0039	0.0622	0	1

Table 6A.1 (continued)

Variable	Mean	Standard Deviation	Minimum	Maximum
College is in CA	0.1388	0.3458	0	1
College is in CO	0.0078	0.0881	0	1
College is in CT	0.0533	0.2246	0	1
College is in DC	0.0260	0.1591	0	1
College is in DE	0.0025	0.0497	0	1
College is in FL	0.0111	0.1047	0	1
College is in GA	0.0169	0.1290	0	1
College is in HI	0.0024	0.0491	0	1
College is in IA	0.0032	0.0561	0	1
College is in ID	0.0009	0.0300	0	1
College is in IL	0.0458	0.2090	0	1
College is in IN	0.0166	0.1278	0	1
College is in KS	0.0014	0.0375	0	1
College is in KY	0.0005	0.0212	0	1
College is in LA	0.0070	0.0836	0	1
College is in MA	0.1339	0.3406	0	1
College is in MD	0.0199	0.1395	0	1
College is in ME	0.0159	0.1250	0	1
College is in MI	0.0173	0.1303	0	1
College is in MN	0.0075	0.0865	0	1
College is in MO	0.0217	0.1456	0	1
College is in MS	0.0007	0.0260	0	1
College is in MT	0.0006	0.0249	0	1
College is in NC	0.0411	0.1986	0	1
College is in NE	0.0012	0.0344	0	1
College is in NH	0.0170	0.1293	0	1
College is in NJ	0.0311	0.1735	0	1
College is in NM	0.0011	0.0327	0	1
College is in NV	0.0005	0.0225	0	1
College is in NY	0.1187	0.3235	0	1
College is in OH	0.0201	0.1405	0	1
College is in OK	0.0011	0.0335	0	1
College is in OR	0.0058	0.0759	0	1
College is in PA	0.0723	0.2589	0	1
College is in RI	0.0320	0.1761	0	1
College is in SC	0.0037	0.0604	0	1
College is in TN	0.0106	0.1023	0	1
College is in TX	0.0185	0.1346	0	1
College is in UT	0.0032	0.0565	0	1
College is in VA	0.0361	0.1866	0	1
College is in VT	0.0110	0.1042	0	1
College is in WA	0.0088	0.0936	0	1
College is in WI	0.0061	0.0781	0	1
College is in WV	0.0001	0.0075	0	1
College is in WY	0.0003	0.0168	0	1

Source: 17,871 college application events among the 3,240 students in the College Admissions Project sample.

Table 6A.2   **Description of the Colleges at Which Students Matriculated, from the College Admission Project Data**

Variable	Mean	Standard Deviation	Minimum	Maximum
Matriculated at this college	1.0000	0.0000	1	1
Admitted to this college	1.0000	0.0000	1	1
Applied early to this college	0.3142	0.4722	0	2
Withdrew application from this college, usually after early decision elsewhere	0.0000	0.0000	0	0
Grants specific to this college	4,029.0040	7,051.1670	0	36,000
Loans from this college	1,020.0040	2,721.6190	0	36,348
Work study amount from this college	296.3472	768.4207	0	15,000
Grant is called a "scholarship"	0.2692	0.4436	0	1
Grant is front-loaded (more in freshman year)	0.0343	0.1820	0	1
Grant is this share of tuition	0.2875	0.5517	0	7
Grant is this share of comprehensive cost	0.1665	0.2728	0	2
Student was a recruited athlete at this college	0.0402	0.1964	0	1
Father is an alumnus of this college	0.0664	0.291	0	1
Mother is an alumna of this college	0.0396	0.1949	0	1
Sibling attended or attends this college	0.0831	0.2761	0	1
College is public	0.2843	0.4512	0	1
College is private not-for-profit	0.7086	0.4562	0	1
College is international, except for Canadian colleges, which are treated as U.S. colleges	0.0068	0.0822	0	1
College's median SAT score, in national percentiles	83.4215	12.5494	32	98
Student's SAT score is this many percentiles *above* college's median SAT score	8.4548	9.1831	0	53
Student's SAT score is this many percentiles *below* college's median SAT score	1.4351	4.8994	0	50
In-state tuition	17,431.8300	9,512.6270	0	27,472
Out-of-state tuition	19,841.1300	6,370.6670	0	27,472
Tuition that applies to this student	18,340.3700	8,599.1560	0	27,472
Room and board at this college	6,821.8120	1,352.4620	0	10,299
In-state comprehensive cost of this college	24,881.0900	10,409.1500	0	35,125
Out-of-state comprehensive cost of this college	27,285.9500	7,335.3150	0	35,125
Comprehensive cost that applies to this student	25,792.1800	9,469.9140	0	35,125
Per-pupil expenditure on students (instruction, student services, academic support, scholarships) of this college, in thousands	29.6174	17.6089	2	78
Instructional per-pupil expenditure of this college, in thousands	19.4170	12.4205	2	72
College is in-state	0.3368	0.4727	0	1
Distance between student's high school and this college, in miles	575.6313	827.2526	0	5,769
College is in AK	0.0000	0.0000	0	0
College is in AL	0.0050	0.0705	0	1
College is in AR	0.0006	0.0250	0	1
College is in AZ	0.0053	0.0727	0	1
College is in CA	0.1199	0.3249	0	1

Table 6A.2 (continued)

Variable	Mean	Standard Deviation	Minimum	Maximum
College is in CO	0.0094	0.0963	0	1
College is in CT	0.0537	0.2255	0	1
College is in DC	0.0265	0.1608	0	1
College is in DE	0.0022	0.0467	0	1
College is in FL	0.0203	0.1410	0	1
College is in GA	0.0131	0.1138	0	1
College is in HI	0.0044	0.0060	0	1
College is in IA	0.0025	0.0499	0	1
College is in ID	0.0022	0.0467	0	1
College is in IL	0.0571	0.2321	0	1
College is in IN	0.0190	0.1367	0	1
College is in KS	0.0025	0.0499	0	1
College is in KY	0.0006	0.0250	0	1
College is in LA	0.0050	0.0705	0	1
College is in MA	0.1218	0.3271	0	1
College is in MD	0.0187	0.1356	0	1
College is in ME	0.0140	0.1177	0	1
College is in MI	0.0194	0.1378	0	1
College is in MN	0.0053	0.0727	0	1
College is in MO	0.0212	0.1442	0	1
College is in MS	0.0012	0.0353	0	1
College is in MT	0.0012	0.0353	0	1
College is in NC	0.0390	0.1937	0	1
College is in NE	0.0022	0.0467	0	1
College is in NH	0.0172	0.1299	0	1
College is in NJ	0.0284	0.1662	0	1
College is in NM	0.0009	0.0306	0	1
College is in NV	0.0022	0.0467	0	1
College is in NY	0.1065	0.3085	0	1
College is in OH	0.0178	0.1322	0	1
College is in OK	0.0022	0.0467	0	1
College is in OR	0.0078	0.0880	0	1
College is in PA	0.0743	0.2623	0	1
College is in RI	0.0300	0.1705	0	1
College is in SC	0.0066	0.0807	0	1
College is in TN	0.0140	0.1177	0	1
College is in TX	0.0225	0.1483	0	1
College is in UT	0.0091	0.0947	0	1
College is in VA	0.0406	0.1974	0	1
College is in VT	0.0106	0.1025	0	1
College is in WA	0.0094	0.0963	0	1
College is in WI	0.0059	0.0768	0	1
College is in WV	0.0000	0.0000	0	1
College is in WY	0.0006	0.0250	0	1

Source: 3,240 college matriculation events among the students in the College Admissions Project sample.

Table 6A.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			
	Low	Medium Low	Medium High	High
Grant is called a "scholarship"	2.584	2.538	2.048	1.313
	(3.11)	(4.39)	(4.26)	(1.24)
Grant is front-loaded (more in freshman year)	1.004	0.971	1.989	1.687
	(0.01)	(–0.06)	(2.00)	(1.28)
Grant is this share of college's tuition	0.760	0.540	1.153	0.621
	(–0.42)	(–1.44)	(0.32)	(–0.75)
Grant is this share of college's comprehensive cost	10.789	24.333	16.695	6.585
	(0.97)	(2.84)	(1.24)	(1.85)

Note: Regressions include all other variables shown in table 6.9.

Table 6A.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		
	Low Selectivity	Medium Selectivity	High Selectivity
Grant is called a "scholarship"	1.927*	1.887*	1.236
	(5.32)	(2.88)*	(0.62)
Grant is front-loaded (more in freshman year)	1.277	1.671	2.599
	(1.00)	(1.00)	(1.44)
Grant is this share of college's tuition	0.789	0.718	2.227
	(–0.80)	(–0.74)	(0.59)
Grant is this share of college's comprehensive cost	27.988*	30.870*	1.315
	(3.08)	(1.87)*	(0.06)

Note: Regressions include all other variables shown in table 6.9
*Statistically significantly different from zero at the 10% level.

Table 6A.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
Grant is called a "scholarship"	1.833	1.934
	(4.69)	(3.98)
Grant is front-loaded (more in freshman year)	1.519	1.130
	(1.74)	(0.29)
Grant is this share of college's tuition	0.782	0.826
	(–0.88)	(–0.37)
Grant is this share of college's comprehensive cost	25.484	20.827
	(2.93)	(1.88)

Note: Regressions include all other variables shown in table 6.9.

References

Avery, Christopher, and Caroline M. Hoxby. 2000. The college admissions project: Counselor report. National Bureau of Economic Research online publication. Available at [http://www.nber.org/~hoxby/collegeadmissions/pdf/counselorreport.pdf].

Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg. 1999. Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings. *Journal of Human Resources* 34 (1): 104–23.

College Board. 2002. Standard research compilation: Undergraduate institutions. Electronic data.

Dale, Stacy, and Alan Krueger. 1999. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. NBER Working Paper no. 7322. Cambridge, Mass.: National Bureau of Economic Research.

Ehrenberg, Ronald, and Daniel Sherman. 1984. Optimal financial aid policies for a selective university. *Journal of Human Resources* 19 (2): 202–230.

Haab, Timothy, and Robert Hicks. 1997. Accounting for choice set endogeneity in random utility models of recreation demand. *Journal of Environmental Economics and Management* 34:127–147.

Hausman, Jerry, and Daniel McFadden. 1984. A specification test for the multinomial logit model. *Econometrica* 52:1219–240.

Hicks, Robert, and Ivar Strand. 2000. The extent of information: Its relevance for random utility models. *Land Economics* 76 (3): 374–385.

Hoxby, Caroline M. 1998. The return to attending a more selective college: 1960 to the present. Harvard University. Unpublished manuscript.

Manski, Charles. 1977. The structure of random utility models. *Theory and Decision* 8:229–254.

McFadden, Daniel. 1987. Regression-based specification tests for the multinomial logit model. *Journal of Econometrics* 34:63–82.

Parsons, George, Andrew Plantinga, and Kevin Boyle. 2000. Narrow choice sets in a random utility model of recreation demand. *Land Economics* 76 (1): 86–99.

Peters, Thomas, Wiktor Adamowicz, and Peter Boxall. 1995. The influence of choice set consideration in modeling the benefits of improved water quality. *Water Resources Research* 613:1781–787.

Peterson's guide to four-year colleges. 2002. Princeton, N.J.: Peterson's.

U.S. Department of Education, National Center for Education Statistics. 2001. *Integrated postsecondary education data system, higher education finance data file.* Electronic data.

———. 2002. *College opportunities online.* Electronic data.

Comment   Michael Rothschild

What a treasure trove this paper is! It is the debut of a new and fascinating data set constructed with skill and sweat. It demonstrates that skilled re-

Michael Rothschild is William Stuart Tod Professor of Economics and Public Affairs, and professor of economics and public affairs at the Woodrow Wilson School, Princeton University, and a research associate of the National Bureau of Economic Research.

searchers can use these data to pose and answer good new questions. Still, it is a discussants job to quibble, and quibble I shall. Simplifying drastically, Avery and Hoxby find the following.

1. Prospective students respond to prices when choosing colleges.
2. In general, the responses are reasonable, but errors are systematic, both in the kinds of errors made and the kind of people who make them.
3. Some students make bad choices.

The authors regard these errors as cause for concern because their sample is restricted to a scarce resource, America's most academically qualified students. A dog that does not bark is the question of what prices in higher education should do. Because their paper is an empirical one, this is no surprise. Still it is worth speculating about equity and efficiency in the allocation of resources in higher education. In the spirit of Avery and Hoxby, assume that each college produces human capital from inputs of student ability and other resources. To fix ideas, suppose that $y_i = R_i^a(G[x_{i1}, \ldots, x_{iN}]^{(1-a)}$, where y_i is output of human capital per capita, R_i is the amount of other resources used per capita, x_{in} is the ability of the nth student attending college i, and $0 < a < 1$. The efficiency problem is to allocate students and resources to colleges so as to maximize total output of human capital. I simplify by assuming both that exactly N students attend each college and that there are K colleges. This is an inessential simplification. The optimal allocation depends on the educational technology. Consider the following three cases:

$$G(x_1, \ldots x_N) = \mathrm{Min}(x_1, \ldots x_N) \quad \text{(MIN)}$$

$$G(x_1, \ldots x_N) = \sum x_n \quad \text{(SUM)}$$

$$G(x_1, \ldots x_N) = \mathrm{Max}(x_1, \ldots x_N) \quad \text{(MAX)}$$

In each case it is easy to characterize the optimal allocation. If the production function is MIN, then students should be allocated so that the most able N students go to the best college, the next N most able students go to the next best college, and so on. Resources are allocated so as to make the labels "best," "next best," and so on descriptive of the allocation of resources. The best college gets a lot of resources, the next best less, and so on. If the production function is MAX, the best student goes to the best college, the next best to the second best, and so on until the Kth best student goes to the worst college. It doesn't matter how the other $N(K-1)$ are allocated to colleges. Allocations of resources are qualitatively similar to the MIN case, but the differences in resources allocation are less pronounced. The SUM case is in between. Students should be allocated to most nearly equalize the sum of ability at each college. Because total ability is nearly equal at each college, other resources should be allocated almost equally to each college. This characterizes efficiency. It is natural to

ask what kind of prices system will support the optimal allocation in each case. I do not think it is difficult to answer this question following the argument in Rothschild and White (1995). I am positive that the efficiency prices do not look anything like the current American price system for undergraduate education.

What about equity? This cannot be answered without specifying some kind of social welfare function. However, some simple observations are possible. Suppose we care about equality of income in the very general sense set out by Atkinson (1970), that is, that we agree that an increase in inequality is something all those who dislike inequality would disapprove of; in other words, an increase in inequality decreases the value of all quasi-concave social welfare functions. Assume that income is determined by human capital and that colleges distribute human capital on a per capita basis. It follows then that the MIN regime is the worst regime and SUM the best; MAX is in between.

Another natural question is what do we know about educational technology in the sense used here. The candid answer is "almost nothing." Surely the technology of the U.S. educational system is not remotely like the technology of any of these examples. However, the allocation of students and resources that we observe in the United States (Winston 1999) bears more than a casual resemblance to the solution to the optimal allocation problem for the MIN technology. Whether this observation justifies this allocation or gives us reason to worry when talented prospective students make the "mistakes" that Avery and Hoxby observe is not clear to me.

References

Atkinson, A. B. 1970. Measurement of inequality. *Journal of Economic Theory* 2 (3): 244–263.

Rothschild, M., and L. J. White. 1995. The analytics of the pricing of higher education and other services in which the customers are inputs. *Journal of Political Economy* 103 (3): 573–586.

Winston, G. C. 1999. Subsidies, hierarchy and peers: The awkward economics of higher education. *Journal of Economic Perspectives* 13 (1): 13–36.