NBER WORKING PAPER SERIES

THE MISSING "ONE-OFFS": THE HIDDEN SUPPLY OF HIGH-ACHIEVING, LOW INCOME STUDENTS

Caroline M. Hoxby
Christopher Avery

Working Paper 18586
http://www.nber.org/papers/w18586

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2012

For inspiration, generous help with data, and numerous useful suggestions, we thank the ACT and The College Board. These two organizations' dedication to providing students with well-informed college application advice is the reason that this research exists. We have been especially helped by Connie Betterton, Michael Matthews, Anne Sturvenant, Ryan Williams, Robert Ziomek. For useful comments and suggestions, we thank Sarah Turner, Eric Bettinger, Scott Carrell, Charles Clotfelter, Susan Dynarski, Bridget Long, Bruce Sacerdote, Douglas Staiger, and Jacob Vigdor. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students
Caroline M. Hoxby and Christopher Avery
NBER Working Paper No. 18586
December 2012
JEL No. I21,I23,I24

ABSTRACT

We show that the vast majority of very high-achieving students who are low-income do not apply to any selective college or university. This is despite the fact that selective institutions would often cost them less, owing to generous financial aid, than the resource-poor two-year and non-selective four-year institutions to which they actually apply. Moreover, high-achieving, low-income students who do apply to selective institutions are admitted and graduate at high rates. We demonstrate that these low-income students' application behavior differs greatly from that of their high-income counterparts who have similar achievement. The latter group generally follows the advice to apply to a few "par" colleges, a few "reach" colleges, and a couple of "safety" schools. We separate the low-income, high-achieving students into those whose application behavior is similar to that of their high-income counterparts ("achievement-typical" behavior) and those whose apply to no selective institutions ("income-typical" behavior). We show that income-typical students do not come from families or neighborhoods that are more disadvantaged than those of achievement-typical students. However, in contrast to the achievement-typical students, the income-typical students come from districts too small to support selective public high schools, are not in a critical mass of fellow high achievers, and are unlikely to encounter a teacher or schoolmate from an older cohort who attended a selective college. We demonstrate that widely-used policies–college admissions staff recruiting, college campus visits, college access programs–are likely to be ineffective with income-typical students, and we suggest policies that will be effective must depend less on geographic concentration of high achievers.

Caroline M. Hoxby
Department of Economics
Stanford University
Landau Building, 579 Serra Mall
Stanford, CA 94305
and NBER
choxby@stanford.edu

Christopher Avery
Harvard Kennedy School of Government
79 JFK Street
Cambridge, MA 02138
and NBER
christopher_avery@hks.harvard.edu

1 Introduction

In this study, we show that a large number--probably the vast majority--of very high-achieving students from low-income families do not apply to a selective college or university.[1] This is in contrast to students with the same test scores and grades who come from high-income backgrounds: they are overwhelmingly likely to apply to a college whose median student has achievement much like their own. This gap is puzzling because the subset of high-achieving, low-income students who do apply to selective institutions are just as likely to enroll and progress toward a degree at the same pace as high-income students with equivalent test scores and grades.[2] Added to the puzzle is the fact that very selective institutions not only offer students much richer instructional, extracurricular, and other resources, they also offer high-achieving, low-income students so much financial aid that the students would often pay *less* to attend a selective institution than the far less selective or non-selective post-secondary institutions that most of them do attend.

We attempt to unravel this puzzle by characterizing low-income, very high-achieving students in the U.S. using a rich array of data, including individual-level data on every student who takes one of the two college assessments, the ACT and the SAT. We divide the low-income, very high-achieving students into those who apply similarly to their high-income counterparts ("achievement-typical" behavior) and those who apply in a very dissimilar manner ("income-typical" behavior). We do this because we are interested in why low-income, high-achieving students appear to base their college-going on their achievement and others base it on their income. We find that income-typical students are fairly isolated from other high achievers, both in terms of geography and in terms of the high schools they attend. In fact, their lack of concentration is such that many traditional strategies

[1] Hereafter, "low income" and "high income" mean, respectively, the bottom and top quartiles of the income distribution for families with a child who is a high school senior. "High-achieving" refers to a student who scores at or the 90th percentile on the ACT comprehensive or the SAT I (math and verbal) and who has a high school grade point average of A- or above. This is approximately 4 percent of U.S. high school students. When we say "selective college" in a generic way, we refer to colleges and universities that are in categories from Very Competitive Plus to Most Competitive in Barron's *Profiles of American Colleges*. There were 236 such colleges in the 2008 edition. Together, they have enrollment equal to 2.8 times the number of students who scored at or above the 90th percentile on the ACT or SAT I. Later, we are much more specific about colleges' selectivity: we define schools that are "reach," "match," and "safety" for an individual student, based on a comparison between his college aptitude test scores and the median aptitude test scores of enrolled students at the school.

[2] Below, we demonstrate that this is true using evidence from the National Student Clearinghouse.

for informing high-achieving students about college--for instance, college admission staff visiting high schools or after-school programs that provide mentoring--are probably prohibitively expensive. We also show that income-typical have only a negligible probability of meeting a teacher, high school counselor, or schoolmate from an older cohort who herself attended a selective college.

In contrast, we show that achievement-typical are *highly* concentrated. Many of these low-income students attend a small number of high schools that admit students on the basis of an exam or previous grades. Since these high schools are nearly all located in large metropolitan areas (not even in smaller metropolitan areas), their students are far from representative of high-achieving, low-income students in general. Moreover, we show evidence that suggests that these schools may be "tapped out"--that is, their students are already so recruited by selective colleges that further recruitment may merely shift students among similar, selective colleges, not cause students to change their college-going behavior in more fundamental ways.

The evidence that we present is descriptive, not causal. This is an important distinction. For instance, we cannot assert that a high-achieving, low-income student would act like an achievement-typical student rather than an income-typical student if he were moved to a large metropolitan area with a high school that practices selective admission. Moreover, we do not assert that income-typical students would have higher welfare if they applied to college in the same way that achievement-typical and high-income, high-achieving students do. We leave such causal tests for related studies in which we are conducting randomized, controlled interventions. Nevertheless, our descriptive evidence makes three important contributions. First, it documents that the number of low-income, high-achieving students is much greater than college admissions staff generally believe. Since admissions staff see only students who apply, it is logical that they underestimate the number who exist. Second, our evidence suggests hypotheses for why so many low-income, high-achieving students apply to colleges in a manner that may not be in their best interests and is certainly different from what similarly high-achieving, high-income students do. Most of the hypotheses are related to the idea that income-typical students--despite being intelligent, literate, and on colleges' search lists (that is, the lists to which selective colleges mail brochures)--lack information or encouragement that achievement-typical students have because they are part of local, critical masses of high-achievers. Third, our descriptive evidence allows us to explain why some traditional

interventions are unlikely to change the situation and allows us to identify other interventions that could plausibly do so.

Our previous work (Avery, Hoxby, Jackson, Burek, Pope, Raman 2006) was perhaps the first to identify the phenomena described in this paper, but there is now a small literature on the topic of "undermatching." We especially note the work of Bowen, Chingos, and McPherson (2009), Dillon and Smith (2012), and Pallais (2009). Relative to those studies, our study's strengths are its comprehensiveness (we analyze the entire population of high-achieving students, not a sample); our complete characterization of each U.S. high school, including its college-going history; our ability to map students to their exact high schools and neighborhoods (this allows us to investigate exactly what they experience); and our use of accurate administrative data to identify students' aptitude, application behavior, college enrollment, and on-time degree completion. The sheer comprehensiveness and accuracy of our data is what allow us to form strong hypotheses about why some high-achieving, low-income students are income-typical and others are achievement-typical. Our data also allow us to assess which interventions might plausibly (and cost effectively) alter such behavior.

The remainder of this paper is organized as follows. In the next section, we present some background on college policies directed toward low-income, high-achieving students. In section 3, we describe our data sources. In section 4, we present a descriptive portrait of very high-achieving U.S. students--their family incomes, parents' education, race, ethnicity, and geography. In section 5, we show that high-achieving students' college application behavior differs greatly by family income. We also show that, conditional on applying to a college, students' enrollment, college grades, and degree receipt do *not* differ by family income (among students with similar incoming qualifications). In section 6, we divide low-income, high-achieving students into achievement-typical and income-typical groups. We then compare factors that might affect the application behavior of the groups. In section 7, we consider several interventions commonly directed towards low-income, high-achieving students, and we demonstrate that they are likely to be cost-prohibitive for income-typical students. We conclude by speculating on the sort of interventions that could plausibly improve income-typical students' welfare.

2 Background on College Policies Directed Toward Low-Income, High-Achieving Students

Many students from low-income families have poor college outcomes: they do not attend college; they drop out before attaining a degree; they earn so few credits each term that they cannot graduate even in 1.5 times the "correct" time to degree; they attend institutions with such poor resources that, even when they do graduate, they earn much less than the median college graduate. These poor college outcomes are often attributed to low-income students being less academically prepared for college and less able to pay for college. These are certainly valid concerns. As we show later, high-income (top income quartile) students are in fact much more likely to be high achievers at the end of high school than are low-income students. Nevertheless, some low-income students *are* very high achievers: at the end of high school, they have grades and college aptitude test scores that put them in the top 4 percent of all U.S. secondary school students or--equivalently--the top 10 percent of students who take one the ACT or SAT college assessment exams.

High-achieving, low-income students are considered very desirable by selective colleges, private and public, which are eager to make their student bodies socio-economically diverse without enrolling students who are unprepared for their demanding curricula.[3] Therefore, these students tend not only to be offered admission by selective schools if they apply, they also tend to be offered very generous financial aid. In recent years, selective schools' aid for low-income, high-achieving students has become so generous that such students' out-of-pocket costs of attendance are sometimes zero and nearly always small.

Figure 1 shows the income distribution for families with a child in the twelfth grade--a good indicator for a family having a child of college-going age in the next year.[4] It demonstrates that the

[3] There are many reasons for selective institutions to prefer socio-economic diversity. These include, to name just a few: a deep respect for merit regardless of need; a recognition that students whose lives were transformed by highly-aided college education tend to be great donors if they do become rich; a belief that a diverse student body makes instruction and research more productive; pressure from politicians.

[4] We pick age 17 because children are nearly always still dependents at that age. If we were to pick age 19, say, we would find numerous families headed by the 19-year-old himself and such families would generally have lower income than the families from which they would have applied to college as "dependent" students. It is such dependent students with whom we are concerned since a negligible share of "mature", independent students are high-

20th percentile of this income distribution was $35,185. Table 1 shows what out-of-pocket costs and loans such a student would have experienced in the 2008-09 school year at a variety of selective and non-selective institutions. The table is organized based on institutions' selectivity as classified by Barrons' *Profiles of American Colleges*: most selective, very selective, selective 4-year institutions, non-selective 4-year institutions, and (non-selective by definition) community colleges and other 2-year institutions. Table 1 also shows the colleges' comprehensive cost for a student who needs *no* financial aid (the "sticker price") and shows their instructional and student-oriented expenditure per student.[5] What the table reveals is that a low-income student who can gain admission to one of the most selective colleges in the U.S. can expect to pay *less* to attend a very selective college with maximum student-oriented expenditures than to attend a non-selective 4-year college or 2-year institution. In short, the table demonstrates the strong financial commitment that selective colleges have made towards becoming affordable to low-income students.[6]

In related work (Avery, Hoxby, Jackson, Burek, Pope, Raman 2006), we analyze Harvard's introduction of zero costs for students with incomes of $40,000 and below starting in 2005. (Harvard is a relevant option for the students we analyze in this paper.) Harvard's policy was quickly imitated or outdone by the institutions with which it most competes: Yale, Princeton, Stanford, and so on. All such institutions subsequently raised the bar on what they considered to be a low enough income to merit zero costs. Thus, even students from families with income above the U.S. median can often attend such institutions for free. Although less well-endowed institutions followed suit to a lesser extent (usually by setting the bar for zero costs at a lower family income than the aforementioned institutions did), the result was very low costs for low-income students at selective institutions, as shown in Table 1.

In our other work, we show that Harvard's policy change had very little effect--at least, very little short-term effect--on the income composition of its class. We estimate that it increased the number

---

achieving according to the criteria we employ, which mimic admissions standards at selective colleges.

[5] Student-oriented expenditure consists of instructional spending, academic support, and student services.

[6] It is interesting to note that a student's out-of-pocket costs and loans, in absolute terms, peak at private colleges of middling to low selectivity. This is because these colleges have little endowment with which to subsidize low-income students and they also receive no funding from their state government (as public colleges do) with which to subsidize students.

of low-income students by approximately 15, in a class of more than 1600. Interestingly, this very modest effect was not a surprise to many college admissions staff. They explained that there was a small pool of low-income, high-achieving students who were already "fully tapped" so that additional aid and recruiting could do little except shift them among institutions that were fairly similar. Put another way, they believed that the *overall* pool of high-achieving, low-income students was inelastic with respect to additional aid and recruiting. While they believed that they might diversify their student bodies by poaching from other selective schools or lowering their admissions standards for low-income students, they did not expect additional aid alone to affect matters much.[7]

In this paper, we show that--viewed one way--the admissions staff are correct. The pool of high-achieving, low-income students who apply to selective colleges is small: for every high-achieving, low-income student who applies, there are about 15 high-achieving, high-income students who apply. Viewed another way, the admissions staff are too pessimistic: the vast majority of high-achieving, low-income students do not apply to *any* selective college. There are, in fact, only about 2.5 very high-achieving, high-income students for every high-achieving, low-income student in the population. The problem is that most high-achieving, low-income students do not apply to any selective college so they are invisible to admissions staff. Moreover, we will show that they are unlikely to come to the attention of admissions staff through traditional recruiting channels.

High-achieving students are traditionally recruited by selective colleges through a few means. First, colleges can purchase from the ACT or College Board mailing lists of students who satisfy certain criteria--for instance, students with high scores on the college assessment exams. Colleges can then send brochures to the student on the mailing lists. However, with rare exceptions, most of these mailings are not differentiated by a student's economic disadvantage.[8] Thus, a low-income,

---

[7] Personal communication with the authors at conferences of the Association of Black Admissions and Financial Aid Officers of the Ivy League and Sister Schools (ABAFAOILSS). Because many very selective institutions practice need-blind admissions, they maintain "Chinese Walls" between the admissions and financial aid offices. Thus, many schools can only *precisely* identify economically disadvantaged students once they have been admitted. However, many admissions officers say that they use students' essays, teachers' letters, parents' education, attendance at an "under-resourced" high school, and similar indicators to identify, provide favorable terms of admission to, and strongly recruit students whom they believe to be economically disadvantaged.

[8] The authors have worked with the ACT, College Board, and a variety of other organizations to create mailing lists that do take account of a student's family income. However, these projects could more accurately be described as pilots than as widespread practice. The main reason that differentiated mailings have not been used is

high-achieving student is likely to receive numerous brochures from selective colleges, but the brochures will give him the same information about the college's costs and financial aid as they would give to a rich student who needed no aid at all. Second, colleges maintain strong relationships with counselors at high schools that reliably send them numerous qualified, low-income applicants. Many of these high schools are those that practice selective admissions, are magnet high schools, or are otherwise well-known "feeders". Third, colleges maintain strong relationships with (and in cases pay for the services of) college access programs that reliably send them qualified, low-income applicants. Since the vast majority of college access programs rely on students to self-select into their activities, it is unclear whether they identify students who would otherwise be unknown to colleges or merely serve as a channel for students to identify themselves as good college prospects. Fourth, college staff visit high schools to describe their offerings and to meet prospective students. However, given that there are more than 42,000 high schools in the U.S., staff can visit only a tiny fraction of schools. A college whose staff visit 100 high schools is considered to be exceptionally dedicated and well-funded. Staff tend to visit schools that are feeders or that are located such that students from numerous high schools can attend the visit. Finally, most colleges welcome on-campus visits from prospective students and often issue special invitations to students from local high schools. Thus, low-income students who happen to live within easy distance of a college campus are quite likely to have opportunities to visit it, although not to visit colleges in general.

3 Data and Identifying High-Achieving, Low-Income Students

We attempt to identify the vast majority of U.S. students who are very high-achieving. Specifically, we are interested in students who are well-prepared for college and who are very likely to be admitted to the majority of selective institutions (if they apply). Thus, as mentioned above, we choose students whose college aptitude test scores place them in the top 10 percent of test takers

---

that the organizations did not know (at least, until their involvement with researchers like the authors) how to identify low-income students with reasonable accuracy. Some colleges do target extra mailings to students who live in low-income zipcodes, but--unlike Census Blocks and Block Groups--zipcodes are not configured to maximize socio-economic uniformity or recognizable neighborhoods. Thus, the zipcode-based mailing lists are very imperfect way of targeting low-income students.

based on either the SAT I (combined math and verbal) or the ACT (comprehensive).[9] Since only about 40 percent of U.S. secondary school students take a college aptitude test, these students are in the top 4 percent of U.S. students. We include in our target group only those students who self-report a grade point average of A- or higher in high school. In practice, this criterion for inclusion hardly matters once we condition on having test scores in the top 10 percent.[10]

Our key data comes from the ACT and the College Board, both of whom supplied us with student-level data on everyone in the high school graduating class of 2008 who took either the ACT or the SAT I.[11] Apart from students' test score history, these datasets contain students' high school identifiers, their self-reported grades, their race and ethnicity, and their gender. Validation exercises have shown that students self-report their grades quite accurately to the ACT and College Board (with just a hint of upward bias), probably because students perceive the organizations as playing a semi-official role in the college application process (Freeberg, 1988). The data also contain answers to numerous questions about students' high school activities and their plans for college.

Importantly, the ACT and College Board data contain a full list of the colleges to which students have sent their test scores. Except in rare circumstances, a student cannot complete his application to a selective college without having the ACT or the College Board send his verified test scores to the college. Thus, score sending is necessary but not sufficient for a completed application. Put another way, score sending may exaggerate but cannot understate the set of selective colleges to which a student applies. Past studies have found that score sending corresponds closely with actual

---

[9] The cut-off is 1300 for combined mathematics and verbal ("Critical Reading") scores on the SAT. The cut-off is 29 for the ACT composite score.

[10] We also considered excluding students who had taken no subject tests since some selective colleges require them. Subject tests include SAT II tests, Advanced Placement (AP) tests, and International Baccalaureate tests. However, we dropped this criterion for a few reasons. First, many selective colleges do not require subject tests or make admissions offers conditional on a student taking subject tests and passing them. Second, among SAT I takers, few students were excluded by this criterion. Third, ACT comprehensive takers usually take subject tests offered by the College Board or International Baccalaureate. When we attempt to match students between these data sources, errors occur so that at least some of the exclusions were false.

We match students between the ACT comprehensive and the SAT I to ensure that we do not double-count high-achieving students. However, this match is easier than matching the ACT comprehensive takers to College Board subject tests, which students often take as sophomores or juniors in high school.

[11] There are approximately 2,400,000 students per cohort who take a College Board test. There are approximately 933,000 students per cohort who take the ACT.

applications to selective colleges (Card and Krueger 2005, Avery and Hoxby 2004). Students who are admitted via an Early Decision or Early Action program often do not apply to colleges other than the one that admitted them early. However, such students typically send scores to all of the schools to which they would have applied had the Early school not admitted them (Avery, Glickman, Hoxby, Metrick, forthcoming). Thus, it is somewhat *better* to observe score sending than actual applications: score sending more accurately reveals the set of selective colleges to which the student would have applied. Note, however, that as most two-year colleges and some non-selective colleges do not require verified ACT or SAT I scores, we do not assume that a student who sends no scores is applying to no postsecondary institutions. Rather, he is applying to no *selective* institution.

For some of our analyses, we need to know where students actually enrolled and whether they are on-track to attain a degree on time (June of 2012 for baccalaureate degrees). We therefore match students to their records at the National Student Clearinghouse, which tracks enrollment and degree receipt. We match all low-income, high-achieving students and 25 percent random sample of high-income, high-achieving students. We do not match all students simply because of expense.

The addresses in the data are geocoded for us at the census block level, the smallest level of Census geography (22 households on average). We match each student to a rich description of his neighborhood. The neighborhood's racial composition, gender composition, age composition, and population density are matched at the block level. Numerous socio-demographic variables are matched at the block group level (556 households on average): several moments of the family income distribution, adults' educational attainment, employment, the occupational distribution, several moments of the house value distribution, and so on. We also merge in income data from the Internal Revenue Service at the zipcode level.

In addition to the data on the graduating class of 2008, we have parallel data for previous cohorts of students who took an ACT or a College Board test. (We have one previous cohort for the ACT; we have more than 10 previous cohorts for the College Board.) We use the previous cohort data in a few ways that will become clear below.

We create a profile of every high school, public and private, in the U.S. using administrative data on enrollment, graduates, basic school characteristics and socio-demographics. The sources are the Common Core of Data (NCES 2009) and the Private School Survey (NCES 2009). By

summarizing our previous cohort data at the high school level, we also create profiles of each school's usual test scores, application behavior, and college plans. For instance, we know how many students from the high school typically apply to each selective college or to any given group of selective colleges. Finally, we add high schools' test scores, at the subgroup level, for each state's statewide test, as mandated by No Child Left Behind. These scores are all standardized to have a zero mean and a standard deviation of one.

We estimate a student's family income rather than relying on the students' self-reported family income. This is for a few reasons. First, both the ACT's and the College Board's family income questions provide a series of somewhat wide income "bins" as potential answers. Second, although the College Board's questionnaire appears to elicit unbiased self-reports of family income, students make substantial unsystematic mistakes when their data are compared to their verified data used in financial aid calculations (the CSS Profile data). Also, about 62 percent of students simply do not answer the College Board's family income question. Third, although the ACT's questionnaire elicits a high response rate, its question refers to the fact that colleges offer more generous financial aid to students with lower family incomes. This framing apparently induces students to underestimate their family incomes: we find that students often report family incomes that are lower than the 10th percentile of family income in their Census block group.

We predict students' family income using the all the data we have on previous cohorts of College Board students, matched to their CSS Profile records (data used by financial aid officers to compute grants and loans). That is, using previous cohorts, we regress accurate administrative data on family income using all of our Census variables, the IRS income variables, the high school profile variables, and the student's own race and ethnicity. In practice, the income variables from the Census have the most explanatory power. However, our goal is simply to maximize explanatory power and many of the variables we include are somewhat multicollinear. We choose predicted income cut-offs to minimize Type I error (false positives) in declaring a student to be low-income. Specifically, we choose cut-offs such that, in previous cohorts, only 5 percent of students who are *not* actually in the bottom quartile of the income distribution are predicted to be "low-income." We recognize that by minimizing Type I error, we expand Type II error, but it is less worrisome for our exercise if we mistakenly classify a low-income student as middle-income than if we do the reverse. This is

because we wish to characterize the college-going behavior of students who are low-income. Since we also find that there are more high-achieving, low-income students than college admissions staff typically believe, we make decisions that will understate rather than overstate the low-income, high-achieving population.

More generally, it is not important for our exercise that our measure of income be precise. What matters for our exercise is that the students we analyze are, in fact, capable of gaining admission and appropriate financial aid at selective colleges. We are confident that they are because, for the most part, we are using the same verified data that the colleges themselves use. Also, we show later that we can accurately predict the colleges at which students enroll, conditioning on the colleges to which they applied. We would not be able to make such accurate predictions if we lacked important achievement and other data that colleges use in their admissions processes.

Hereafter, we describe as low-income any student whose estimated family income is at or below the cut-off for the bottom quartile of the 2008 distribution of incomes among families who had a child in his senior year of high school: $41.472.[12] We describe as high-income any student whose estimated family income is at or above the cut-off for the top quartile of the same distribution: $120,776. See Figure 1 for other percentiles.

4 A Portrait of High-Achieving Students in the U.S.

Who and where are the high-achieving students in the U.S.? In this section, we briefly characterize them, leaving more detailed analysis of the low-income, high-achieving group for later.

Figure 2 shows that 34 percent of high achievers have estimated family income in the top quartile and 27 percent have estimated family income in the third quartile. That is, high income families are overrepresented in the high-achieving population. However, 22 percent and 17 percent of high achievers have estimated family incomes in, respectively, the second and bottom quartiles. We conservatively estimate that there are *at least* 25,000 and probably as many as 35,000 low-income high achievers in the U.S. Recall that our methods tend to understate the number of low-

[12] Since we require microdata to create the relevant distribution, our source for this information is the American Community Survey 2008.

income students.

Table 2 shows that among high achievers, those who are from higher income families do have slightly higher college aptitude scores, but the difference is small. The average low-income high achiever scores at the 94.1th percentile. The average high-income high achiever scores at the 95.7th percentile.

Data on the parental education of high achievers are unfortunately very incomplete because ACT takers are not asked to report their parents' education and 52 percent of SAT I takers fail to answer the question about their parents' education. Moreover, SAT I takers are apparently less likely to report their parents' education when it is low. We base this assessment on the observation that parents' education is more likely to be missing for students who live in Census block groups with low adult education. For what they are worth, however, the data on the parents' education are shown in Figure 3.[13] More precisely, we show the maximum of father's reported educational attainment and mother's reported educational attainment. 50.7 percent of students who report their parents' education say that at least one parent has a graduate degree. 27.9 percent say that at least one parent has a baccalaureate degree and another 6 percent cite "some graduate school" (but no degree). 11.6 percent claim that at least one parent has an associate's degree or "some college or trade school" (but no degree). Only 3.8 percent report having parents neither of whom has more than a high school degree. Perhaps the most interesting thing about the parents' education data is that it seems to indicate that high achievers are reluctant to report that they have poorly educated parents. This is in contrast to family income data from the same College Board questionnaire. Many students did not answer the income question but those who did answered it in an unbiased (albeit fairly inaccurate) way.

Figure 4 displays information on high achievers' race and ethnicity, which 98 percent of students voluntarily report. 75.8 percent of high achievers say that they are white non-Hispanic and another 15.0 percent say that they are Asian. The remaining 9.2 percent of high achievers are associated with

[13] We do not attempt to correct these data for biases because we do not have verified data on parents' education that we could use to estimate the errors accurately. This is in contrast to family incomes, where we do have a source of verified data (the CSS Profile).

an underrepresented minority.[14] They are Hispanic (4.7 percent), black non-Hispanic (1.5 percent), Native American (0.4 percent), or mixed race/ethnicity (2.6 percent). If we focus on *low-income* high achievers only (Figure 5), we see that 15.4 percent are underrepresented minorities. Interestingly, the entire increase in this share comes out of the percentage who are white. Asians make up 15.2 percent of low-income high achievers, almost identical to their share of all high achievers.

A key take-away from Figure 5 is a student's being an underrpresented minority is not a good proxy for his being low-income. Thus, if a college wants its student body to exhibit *income* diversity commensurate with the income diversity among high achievers, it cannot possibly attain this goal simply by recruiting students who are underrepresented minorities. If admissions staff do most of their outreach to low-income students by visiting schools that are largely Hispanic and black, the staff should realize that this strategy is likely to lead to a student body that is not income-diverse.

The choropleth map in Figure 6 shows the number of high achieving students in each county of the U.S. Counties are an imperfect unit of observation because some are large in land area and some are small. Nevertheless, they are most consistent political units in the U.S.[15] The darker is the county's coloring, the more high-achieving students it contains. What the map demonstrates is that critical masses of high-achieving students are most likely to be found in urban counties in southern New England (Massachusetts, Connecticut, Rhode Island), the Mid-Atlantic (New York, New Jersey, eastern Pennsylvania), southern Florida, and coastal California from Bay Area to San Diego. The other critical masses are more scattered, but a person familiar with U.S. geography can pick out Chicago (especially), Houston, Dallas/Fort Worth, Atlanta, and some smaller cities. In short, if one's goal were to visit every county where one could gather at least 100 high achievers, one could concentrate entirely on a limited number of cities on the east and west coasts and a few cities in between.

Some part of the above statement is due to the fact that high-income, highly educated parents are somewhat concentrated in the aforementioned areas and such parents, we have seen, are

[14] Underrepresented minority is the term of art in college admissions. Notably, it excludes Asians.

[15] That is, the size and scope of municipalities, school districts, and other jurisdictions are far *less* consistent than those of counties.

somewhat more likely to have high-achieving children. However, some part of the above statement is due purely to population density. That is, even if children in all counties of the U.S. were equally to be high-achieving, there would still be critical masses of them in densely populated counties and *vice versa*. The choropleth map in Figure 7 illustrates the role of population density by showing the number of high-achieving students *per 17 year old* in each county. The darker a county is, the higher is its decile on this relative measure. The map makes it clear that this relative measure is far less concentrated than the absolute measure that favors dense counties. In fact, there is a belt of counties that tend to produce high achievers that runs from Minnesota and the Dakotas south through Missouri and Kansas. There are also a good number of Appalachian, Indiana, and non-coastal California but still Western counties that tend to produce high achievers. In short, if one's goal were to meet a representative sample of high achievers, one's trip could not be concentrated on a limited number of counties on the Coasts and a few cities in-between.

5 College Applications, Enrollment, and Degree Receipt among High-Achieving Students in the U.S.

In this section, we analyze the college application choices, enrollment decisions, and on-time degree receipt of high-achieving students in the U.S., paying attention to how low-income students behavior does or does not differ from that of high-income students. Because colleges in the U.S. are so varied and large in number, we will characterize them by the college aptitude score of their median student, expressed in percentiles of the national college aptitude test score distribution. This statistic, although admittedly insufficient to describe them fully, has important qualities. First, it is probably the single best, simple indicator of selectivity--much better the admissions rate, for instance (Avery, Glickman, Hoxby, and Metrick, forthcoming). Second, when an expert college counselor advises students on how to choose a portfolio of schools to which to apply, he usually tells students to apply to a few schools that are a "reach," four or more schools that are "par" or "match," and one of more schools that are "safe." Similar advice is widely available on internet sites of college advising organizations with a strong reputation, including the College Board and the ACT. Experts use schools' median scores to define "reach" (typically: the school's median score is five or more

percentiles above the student's own), "match" (typically: the absolute value of the difference between the school's median score and the student's own is less than five percentiles, and "safety" (typically: the school's median score is five to fifteen percentiles below the student's own). Naturally, the exact cut-offs for these categories vary by expert, and high-achieving students are often advised to apply to their state's public flagship university, even if it falls below the safety zone.[16] High-achieving students are generally advised to apply to at least eight schools.

a College Application Behavior -- A Graphical Analysis

In this sub-section, we provide graphical evidence of what student's application portfolios look like. This is somewhat informal but useful for fixing ideas and defining categories before we move to formal econometric analysis in the next sub-section. In what follows, an "application" is sending a score to a college.[17]

Figure 8 shows a histogram of the application portfolios of high income students. It is important to understand how this figure is constructed. On the horizontal axis is the difference between the applied-to college's median college aptitude score and the student's own score, in percentiles. Thus, if an application is located at 0, the student is applying to a match school whose median student has exactly his scores. An application at +8 is a reach. An application at, say, -13 is a safety. Since non-selective colleges do not require their students to take college aptitude tests (and thus do not report a median student score), an application to a non-selective school is placed at -94, which is 0 minus the average percentile score of high-achieving students in the data. It is not obvious where to place applications to non-selective schools, but -94 has the advantage that such applications cannot be mistaken for applications to a school that is selective but that sets a very low bar.

---

[16] Experts also advise students to look at the high school grade point average that is typical of a college's student. However, such grade-based categories are not terribly relevant to high-achieving students because selective colleges vary so much more on the basis of college aptitude test scores than on the basis of high school grades.

[17] As noted above, a student may often apply to a non-selective college without sending scores, although a good number of students send scores to them for apparently no reason (the first sends are free) or for placement purposes (that is, to avoid being placed in lower-level or even remedial courses). If we match students to their enrollment records in the National Student Clearinghouse, we can add to their set of applications any non-selective school in which they enrolled without sending scores. This does not change the figures much although it does systemically amplify the bar for non-selective applications. We do add applications in way for the analysis in the second half of this section, but it makes too little difference here to be worthwhile--especially as we would then have to show figures for sample of the students, rather than the population of them.

Each student is given a weight of one in the histogram and this weight is split evenly over his applications. This is to ensure that the histogram does not over-represent the behavior of students who apply to more schools since, after all, each student will only enroll at a single one (initially). Thus, if a student puts all of his eggs in one basket and applies to a single +8 school, his full weight of one will show up in the +8 bar. If a student applies to one +8 school, one +7 school, and so on down to one -8 school, one 17th of his weight will show up in each of the relevant bars. Note that each bar is 2 percentiles wide.

Figure 8 shows that high income students largely follow the advice of expert counselors. The bulk of their applications are made to match schools. They apply to some reach schools as well, but they are mechanically limited in the extent to which they can do this. There are *no* reach schools for slightly more than half of the high-achieving students we study.[18] High-income, high-achieving students also apply fairly frequently to safety schools. It is noteworthy that many such students apply to their state's flagship university, probably as a safe school but possibly out of state loyalty. These schools vary greatly in selectivity, so that some such applications are in the safe range, but other applications to flagships appear far more safe than anyone would think necessary. For instance, an application to a flagship with median scores at the 50th percentile would end up at -40 to -50.

The reader might be surprised to find that high income, high-achieving students apply to some colleges that are non-selective on academic grounds. However, these appear to be students with special abilities and tastes because the schools in question are often specialty schools--music conservatories, art or design schools, drama or performing arts schools, cooking schools, and so on.

Figure 9 shows that few low-income, high-achieving students follow the advice of expert counselors. More than 40 percent of the mass in the histogram loads on non-selective schools. (This is an underestimate because scores are not sent to some non-selective schools. If we include every non-selective enrollment as a non-selective application, the non-selective bar on the histogram would rise by 5.1 percent.) Moreover, the non-selective colleges to which low-income students apply are rarely of the performing arts type mentioned above. They are often local community colleges or

[18] For instance, consider a student whose own scores put him at the 94th percentile. In order to apply to a reach school, he would need to apply to a school whose median student scored at the 99th percentile. There are no such schools--at least no schools that admit to having such a high median score.

local four-year institutions with meager resources per student and low graduation rates. Much of the size of the non-selective bar is due to the fact that many low-income, high-achieving students apply only to a single non-selective college or to a single non-selective college and one additional college, which is often only weakly selective. In fact, 53 percent of low-income, high-achieving students fit the profile we will hereafter describe as income-typical: they apply to no school whose median score is within 15 percentiles of their own and they do apply to at least one non-selective college.

At the other extreme, 8 percent of low-income, high-achieving students apply in a manner that is somewhat close to what is recommended and to what their high-income counterparts do: they apply to at least one match college, at least one safety college with median scores not more than 15 percentiles lower than their own, and apply to no non-selective colleges. We hereafter designate such students as achievement-typical, noting once a student fits the above criteria, he usually also applies to other match colleges.

The remaining 39 percent of low-income, high-achieving student use application strategies that an expert would probably regard as odd. It is not unusual, for instance, to see students who apply to only a local non-selective college and one extremely selective and well-known college--Harvard, for instance. No expert would advise such a strategy because the probability of getting into an extremely selective, well-known college is low if a student applies to just one--even if the student's test scores and grades are typical of the college's students. Another strategy that appears fairly often is for a student to apply to a single public college within his state that is selective but is much less selective than the state's flagship university. Although about half of these application choices could be motivated by distance from home, the other half cannot because the flagship university is nearer. Another strategy that falls into the idiosyncratic category is a student applying to a single private college outside his state that *is* selective but much less selective and much poorer in resources than the student's private match colleges would be. Such choices are odd because although the private match colleges might offer fewer scholarships that are explicitly merit-based, they offer much more generous need-based aid so that the student would pay less to attend and would enjoy substantially more resources. Furthermore, it is almost never sensible for a low-income student to apply to a single private, selective college: he can use competing aid offers to improve the aid package he gets

from his most preferred college.

We have described a few salient strategies that appear among low-income, high-achieving students who are neither achievement-typical nor income-typical. However, most of these students do not fit an obvious pattern. Thus, below we turn to an econometric analysis, in which we can simultaneously consider a large number of factors correlated with students' application choices.

Before doing this, however, we offer Figure 10, which overlays the histograms for low-income, middle-income, and high-income students who are high-achieving. It cuts off the portion of the histogram that shows non-selective colleges so that readers can focus on application choices among colleges that selective to at least some degree. It will be observed that middle-income students' behavior is about midway between that of their low- and high-income counterparts. Moreover, even within the subset of applications that are made to selective colleges, high-income students apply much more to match colleges and low-income students apply much more to colleges far below the safety level.

b College Application Behavior -- An Econometric Analysis

In this sub-section, we assess the factors that are associated with a student's choice of his application portfolio using the conditional logit model in which a student can apply to all colleges in the U.S. but decides to apply only to some. This model is based on a random utility framework and assumes that the student prefers all colleges to which he applies over the colleges to which he does not apply. We do not assume anything about the student's preference ordering within the colleges to which he applied.[19] Each possible college matched with each student is an observation and the dependent variable is a binary variable equal to one if the student submits an application to the college and zero otherwise.

The explanatory variables we consider are the difference between a school's median test score

[19] We considered estimating a rank-ordered logit model (Beggs, Cardell, and Hausman 1981) based on the assumption that the order in which the student sent scores to colleges indicates the rank order of his preference among them. (All colleges to which no application are assumed to generate net utility below the bottom ranked college.) If we do this, the rank-ordered logit generates fairly similar results, in part because many students do not send scores to more than a few colleges. However, the order of score sending might be a poor proxy for some students' preference orderings because they choose a first batch of colleges to receive their scores before they know what those scores are. Once they learn their scores, they choose a second batch of colleges to receive their scores. At application time, they presumably prefer the second batch to the first.

and the student's own test score if positive, the same difference if negative,[20] an indicator for the school's being non-selective, the distance between the student's home and the school, the square of this distance, an indicator for the school being the most proximate, an indicator for the school being public, an indicator for the school being in-state for the student, an indicator for the school being the flagship university of the student's state of residence, the sticker price of the college, the likely net cost of the college for the student, and the student-oriented resources per student at the college. We fully interact these explanatory variables with indicators for the student being low-income, high-income, or in between. Thus, we estimate separate coefficients for each income group. In the tables, we do not show the coefficients for the middle income group because they nearly always fall between those of the high- and low-income students, but the coefficients are available upon request.

Table 3 shows the results of this estimation. The coefficients are expressed as odd ratios so that a coefficient greater than one means that an increase in the covariate is associated with increase in the probability that the student applies to the school, all other covariates held constant. Based on our graphical analysis, we expect to find very different coefficients for low- and high-income students, and we do.[21] Note that, although it is convenient to describe the coefficients as though they literally revealed preference, they should not be given such a strong interpretation or a causal interpretation. For instance, students might "disfavor" distance not because distance itself generates negative utility but because distant schools have, say, distinct cultures that the student dislikes.

We find that high-income students strongly favor reach colleges and disfavor safety colleges (those for which the score difference is negative). Per percentile of difference, this effect is much stronger on the reach side than on the safety side but recall that high-achieving students can only reach a bit whereas they can apply to *very* safe schools. High-income students strongly dislike non-selective institutions. High-income students dislike higher net costs but (all else equal) like higher sticker prices. This is probably because higher sticker prices are associated with higher per-student

[20] That is, we do not assume that the response of a student to mismatch is symmetric around his own test score. He may only slightly like being at a reach school, for instance, but strongly dislike being at a safety school.

[21] In Avery and Hoxby 2004, we found much smaller differences in the behavior of low- and high-income students, but all the students we sampled attended high schools that were at least somewhat reliable feeders. As shall be seen, the low-income students we sampled were thus very disproportionately Recommendation-Takers who *do* behave fairly similarly to high-income students.

resources, one measure of which they also like. High-income students dislike distance, but the quadratic term is such that they dislike it only up to a point, after which they are fairly indifferent. They have a mild preference for in-state schools and their state's flagship university. They do not have a statistically significant preference for publicly controlled schools.

The low-income students exhibit several immediate contrasts. Such students strongly favor non-selective colleges. This was obvious in the graphical evidence. They do not disfavor schools whose median scores are lower than theirs. They slightly disfavor schools with higher sticker prices (recall that these were favored by high-income students) and do not have a preference for net costs that is statistically significantly different from zero. Low-income students do favor schools with higher expenditure per student, but not nearly as much as high-income students do. Distance is strongly disfavored for schools within 100 miles but, thereafter, low-income students are fairly indifferent to it. Low-income students favor in-state schools somewhat more than high-income students do, but low-income students do not exhibit a preference in favor of their state's flagship university. They slightly favor publicly controlled colleges.

Table 4 repeats the estimation but interacts the covariates with indicators for high-income, middle-income, low-income achievement-typical students, low-income income-typical students, and other low-income students. The estimated coefficients for achievement-typical students are fairly similar to those of high-income students. It is the income-typical students whose coefficients are strikingly different. Of course, these results are somewhat by design, given the way we categorized low-income students into achievement-typical and income-typical groups. However, the coefficients validate the categorization: achievement-typical students *do* pursue similar application strategies as high-income students. In the next section of the paper, we assess which factors predict a student being achievement-typical and which predict his being income-typical.

c College Enrollment and Progress toward a Degree

In this sub-section, we demonstrate that, *conditional on applying to a specific college*, high- and low-income students thereafter behave similarly. There is no statistically significant difference in their probability of enrolling or in their progress toward a degree.

To find the first of these results, we estimate a conditional logit model in which the binary outcome is 1 for the college in which the student initially enrolled and 0 for all others. Importantly,

we limit the student's choice set to the colleges to which he applied. So that the student's enrollment decision is compared to those of students who applied to the same college, we include a fixed effect for each college. We also include interactions between these fixed effects and an indicator for a student's having high or low income. We then test whether each college's high-income or low-income interaction is statistically significantly different from zero. Thus, we test, specifically, whether high- and low-income students who apply to the same college are differentially likely to enroll in it.

Because there are so few high-income students who apply to non-selective and low selectivity colleges, many of the high-income-by-college-indicators are dropped by the regression. Therefore, the effects of income on enrolling in such colleges, conditional on having applied, is not identified. This is noted in the table.

The reader should observe that this test subsumes colleges' admissions decisions. That is, if we find that high- and low-income students are equally likely to enroll in a college, conditional on having applied to it, they must be getting treated similarly in the admissions process. Otherwise, they would enroll differentially simply because they had been admitted differentially.[22]

Table 5 shows the results from this estimation. The table is organized by colleges' median test scores, with more selective colleges closer to the top. We find that only very small shares of low-and high-income enrollment probabilities (conditional on applying) are statistically significantly different from one another at the 5 percent level. For instance, low-income enrollment probabilities differ from high-income enrollment probabilities in only 4 percent of the colleges that have median scores at the 90th percentile or above. This is about what we expect from a test at the 5 percent level. The remaining rows of the table contain similar results, all suggesting that low- and high-income students do not enroll differentially, conditional on applying.

Our test for differential progress toward a degree, conditional on the school at which a student initially enrolled, is constructed in an analogous way. The dependent variable is now the percentage of coursework toward a four-year degree that the student appears to have completed as of June

[22] We can interact additional student characteristics that might affect admission--for instance, race and ethnicity--with colleges' fixed effects. This effectively "soaks up" each college's preferential admissions standards. However, such a specification does not change the estimated coefficients of interest to a noticeable extent and it makes interpretation slightly harder.

2012.[23] A student who is making on-time progress should have completed 100 percent of his coursework by then. We estimate a fixed effect for every college so that students are compared to others who enrolled in the same school. We interact the fixed effects with high- and low-income indicators, and we test whether these interactions are statistically significantly different. Again, the effects for non-selective and low selectivity colleges are not identified because so few high-income students enroll in them.

Table 6, which is organized like Table 5, shows the results from this estimation. For selective colleges, we find that only very small shares of colleges have statistically significantly differences between the progress of their low- and high-income students. For instance, low-income students' progress toward a degree differs from high-income students' progress toward a degree at only 5 percent of the colleges that have median scores at the 90th percentile or above. This is what we expect from a test at the 5 percent level.

There are two key take-aways from this sub-section. First, the *application stage* is where interesting differences appear in the behavior of high-income, high-achieving students and low-income, high-achieving students. If they apply to the same colleges, their educational paths are similar afterwards. Thus, interventions that could make low-income, high-achieving students' college careers look more like those of their high-income counterparts must, as a logical matter, be focused on the application stage or preparation for it. Second, the data do not suggest that low-income students who currently fail to apply to selective colleges and therefore fail to attend one would perform badly if they did attend one. Of course, we cannot say that they would perform just as well as the low-income students who do apply and do enroll. We would need to estimate causal effects to make such a claim and we do not attempt to do that in this paper.[24] However, we are certainly not struck by evidence that low-income students have poor outcomes when they apply to selective schools, and it is worth remembering that not all low-income students who do this are achievement-typical. Some are students who are neither achievement-typical nor income-typical but

[23] We do not consider progress toward a two-year degree because virtually none of the high-achieving students reported that a two-year degree was their educational goal in the descriptive questionnaires that accompany the ACT and SAT I tests.

[24] However, Hoxby's ongoing work with Sarah Turner will attempt to estimate these causal effects.

who pursue an idiosyncratic application strategy such as applying to one non-selective college and to Harvard.

6 Factors that Predict a Student's Being Achievement-Typical or Income-Typical

In this section, we use simple descriptive statistics to identify some factors that predict whether a low-income student is achievement-typical or income-typical. Our goal in this section is to characterize the two types of low-income students sufficiently well that we can build hypotheses about why they apply to colleges so differently.

Ex ante, our hypotheses fall into three broad categories:

(i) Despite the fact that both income-typical and achievement-typical students have estimated family incomes in the bottom quartile, income-typical students are found to be socio-economically disadvantaged compared to achievement-typical students when we examine their backgrounds more carefully.

(ii) Income-typical students are likely to be poorly informed about college compared to achievement-typical students.

(iii) Income-typical students are making rational, well-informed choices about college. Their utility from attending non-selective or less selective colleges exceeds the utility they would derive from attending more selective colleges.

We can look for evidence of hypotheses in categories (i) and (ii). The hypothesis in category (iii) is inherently untestable, so it is effectively the residual explanation if there is no evidence for other hypotheses. Note that if hypothesis (iii) is the true one, students need not get more utility from attending a non-selective college because it is a good academic match for them. A student might attend a school that is obviously a poor academic match because it enables him, say, to look after his family. He might derive sufficient utility from doing this so that his college choice is utility maximizing.

Table 7 shows several family factors that might reveal that income-typical students are truly socio-economically disadvantaged relative to achievement-typical students. The statistics tends to

go the wrong way for hypotheses of type (i). Income-typical students have slightly higher estimated family income than achievement-typical students do. The (admittedly very flawed) reports of parents' education suggest that income-typical students' parents might have 0.7 years more of education than those of achievement-typical students. Achievement-typical students are more likely to be black or Hispanic, so they are presumably more, not less, likely to have experienced discrimination or expect to experience it in college.

Table 8 shows several neighborhood factors that are useful for assessing hypotheses of both types (i) and (ii). A person's neighbors reveal something about his own socio-economic disadvantage, but they also reveal something about the information he is likely to encounter. The statistics show that income-typical and achievement-typical students live in Census block groups with very similar family income. However, achievement-typical students' block groups are less white, and more black, Hispanic, and Asian that those of income-typical students. Achievement-typical students also have more baccalaureate degree holders in their block groups--both in absolute number (207 versus 144) and as a share of adults (22.0 percent versus 16.8 percent). This last fact suggests that income-typical students may be less likely to get advice about college from a neighbor with a degree.

Table 9 compares the geography of income-typical and achievement-typical students, and the contrast is striking. 65 percent of achievment-typical students live in the main city of an urban area, whereas only 30 percent of income-typical students do. Even within the main city residents, achievement-typical students are much more likely to live in a large urban area (one with population greater than 250,000). Indeed, 70 percent of the achievement-typical students come from *just fifteen* urban areas: San Francisco, Oakland, Los Angeles, San Diego, Dallas, Houston, Chicago, Cleveland, Pittsburgh, Portland, Boston, Providence, New York, Philadelphia, and Baltimore. Only 21 percent of achievement-typical students live in a non-urban area (not necessarily rural, but a town rather than an urban area suburb). 47 percent of income-typical students live in a non-urban area. Put another way, income-typical students tend to be the high-achievers who live in counties that had a large number of high-achievers *per 17-year-old* (Figure 7) but not a large number of achievers in absolute terms (Figure 6).

Using administrative data from the U.S. Department of Education, Table 10 compares the

schools attended by income-typical and achievement-typical students. The statistics should help us to assess the students' academic disadvantages and also the amount of college-related information they might obtain at school. Achievement-typical students are considerably more likely to attend a school that is classified as a magnet school or an independent (as opposed to religious) private school. These statistics certainly understate the extent to which the achievement-typical students attend high schools that admit students on the basis of exams or grades. Finn and Hockett (2012) find that only a small share of such high schools are classified as magnet schools.[25] The per-pupil spending of achievement-typical students' public schools is higher, but since facilities and staff costs are often higher in the urban areas where they tend to live, it is unclear whether they are actually advantaged by the higher spending. The pupil-teacher and pupil-counselor ratios are fairly similar for achievement-typical and income-typical students: 18.3 versus 17.2 and 328 versus 305.

Using survey data from the Schools and Staffing Surveys from 1987 to 2007 and data on previous cohorts from the College Board and ACT, Table 11 compares college-related factors at the high schools attended by achievement-typical and income-typical students.[26] The first striking statistic in the table is what a tiny share of low-income students' teachers graduated from colleges that would be match or safety colleges for high-achieving students. Only 1.1 percent of income-typical students' teachers attended match colleges and only 5.0 percent attended safety colleges. The shares are larger for achievement-typical students' teachers, but still not large: 2.9 percent from match colleges and 7.5 percent from safety colleges. Even high-income students do not encounter many teachers with degrees from very selective colleges.

Income-typical students attend high schools where just 1.6 percent of students in previous

[25] Finn and Hockett found most of the selective high schools in their study by word of mouth and contacting all high schools that were so dissimilar to other school in their district that they seemed likely to practice selective admissions. Interestingly, many school districts deemphasize the existence of their selective high schools, which can be controversial. This is perhaps why there was no reasonably accurate list of them prior to Finn and Hockett (2012).

[26] We use all of the Schools and Staffing surveys in an attempt to pick up as many high schools as possible, but we nevertheless end up with teacher data for only 34 percent of the high-achieving students we study. We use the survey weights to create statistics that should be nationally representative. For the statistics based on previous cohorts, we use the actual previous cohorts from the College Board but must assume that our one previous cohort from the ACT was representative of the whole previous decade.

cohorts applied to one of the top ten most selective colleges in the U.S.[27] In contrast, 7.6 percent of students so applied from the previous cohorts of achievement-typical students' schools. In addition, only 3.8 percent of income-typical students' high school class consists of high-achievers (including the student himself) whereas 11.2 percent of achievement-typical students' class consists of high-achievers. Since income-typical students' high schools are, on average, less than two-thirds the size of achievement-typical students' high schools, their low percentages translate into very little school-based contact with other high achievers or high school alumni who attended selective colleges. The low percentages also suggest that their counselors are unaccustomed to advising students who have opportunities to attend selective colleges.[28]

Of course, one might gather and advise a critical mass of high achievers outside of the high school setting, but the bottom rows of Table 11 show that even this is difficult for income-typical students. The radius needed to gather 50 high-achievers is 37.3 miles for the average income-typical student, where as it is merely 12.2 miles for the average achievement-typical student. Since a college access program cannot expect to get participation from *every* qualified student in the area it covers, the radii shown suggest that most income-typical students cannot be reached by programs that require a critical mass of high-achievers to operate at efficient scale.[29]

7 Thought Experiments: Interventions that Might Inform Income-Typical Students

[27] Arguably, focusing on these colleges overstates the extent of previous cohorts' sophistication about college applications. These colleges are the most likely to show up in odd strategies like applying to one non-selective institution and to Harvard.

[28] At College Board sessions attended by high school counselors, the authors have received similar suggestions from the counselors themselves. Strikingly, several counselors reported that when the rare student in their school was qualified to attend very selective colleges, they told him to guide himself by gathering information independently on the internet because they lacked expertise. This is despite the fact that counselors who attend College Board sessions are probably more sophisticated and informed than the average counselor.

[29] Of course, there are college access programs in many areas where Undermatchers tend to live, but the typical program is focused on motivating students merely to attend college (versus not attending). Programs tend to focus on the decisions faced by high-achieving students with many college opportunities, and they do not provide much advice on negotiating the multi-layered application process that very selective colleges use.

In this section, we consider a few interventions that might affect how informed income-typical students are about their college-going opportunities. We do this because, as shown in the previous section, the data evince no support for hypothesis i (that income-typical students are actually more disadvantaged than achievement-typical ones) so that we are left with hypotheses ii (students are poorly informed) and iii (students are well-informed and utility-maximizing). One way to assess whether hypothesis ii is to consider what information actually reaches or could reach income-typical students. After all, they are high-achieving, low-income students who are apparently desirable applicants. Why should they not, for instance, become informed in the process of being recruited by colleges?

Colleges often send admissions staff to high schools in order to recruit high-achieving. Therefore, let us consider a thought experiment whereby any student who attends a high school that contains at least 20 high achieving students will have contact with some college admissions staff. (We chose a cut-off of 20 because it is expensive in time and money for admissions staff to visit high schools in which they cannot fill at least a classroom with potential applicants.) If this experiment occurred, 92 percent of high-income, high-achieving students and 66 percent of achievement-typical students would have contact with admissions staff, but only 17 percent of income-typical students would have such contact.

Of course, admissions staff can hold evening or weekend events that students from multiple high schools can attend. Thus, we should also consider what would happen if admissions staff visited every location in the U.S. where they could gather at least 20 high achieving students from a 10 mile radius. Such visits would ensure that 94 percent of high-achieving, high-income students and 73 percent of achievement-typical students could meet with admissions staff. But such visits would allow only 21 percent of income-typical students to meet admissions staff.

Clearly, admissions staff visiting students is unlikely to be an effective method of informing income-typical students. What about students visiting colleges? As a thought experiment, consider what would happen if every high achieving student visited colleges if he could reach five "match" colleges by traveling 2000 miles or less. 75 percent of high-income, high-achieving students and 71 percent of achievement-typical students would do a college "tour." Only 22 percent of income-typical students would.

In fact, remembering that 70 percent of achievement-typical students are drawn from only fifteen urban areas, we note that many of these students need not travel at all to visit one or more selective colleges. In fact, without needing anything other than a subway pass, a New York City student could easily visit Columbia, Barnard, New York University, Cooper Union, and at least six other colleges that are "Very Competitive" or more selective. A Boston, San Francisco Bay Area, Los Angeles, Chicago, or Philadelphia student would also be spoiled for choice. Even a student from Portland, Maine--an area that might have seemed out of place on the list of fifteen urban areas--has Bates, Bowdoin, Colby, and Dartmouth (all very selective institutions) within a modest radius. In fact, we know from colleges' own published materials and communications with the authors that many colleges already make great efforts to seek out low-income students from their metropolitan area. These strategies, while successful to some degree, fall somewhat under the heading of "searching under the lamp-post." That is, many colleges look for low-income students where the *college* is instead of looking for low-income students where the *students* are.

We have already seen that income-typical students are very unlikely to encounter a teacher, counselor or neighbor who attended a selective college, so interventions that depend on their presence are unlikely to be effective. Also, the logic that made admissions staff visits ineffective with income-typical students works similarly for after-school or weekend college access programs: programs with sustainable costs are unlikely to reach income-typical students.

What are some interventions that might inform income-typical students about college and are any of them tried on a large scale, so that their effectiveness might be evaluated? First, a college has many more alumni than admissions staff, and alumni are much more broadly based, geographically, than admissions staff. For instance, the anonymous private, very selective university studied by Meer and Rosen (2012) has at least one alum in nearly every U.S. county.[30] Presumably, it would be possible for college to give their alumni the names of local students who appear on the "search lists" of students who are likely qualified for admission. Such alumni-based information interventions might work well because they are unlikely to be hampered by the lack of geographic concentration among income-typical students. The main challenges for such interventions would

[30] Meer and Rosen generously computed the relevant statistics for the authors.

seem to be the need to coordinate and inform alumni. It would be problematic, for instance, if alumni knew very little about their college's current curriculum or financial aid policies.

Income-typical students are intelligent and able to absorb written material. Thus, other interventions that might affect them would be purely informational, written interventions--distributed by mail, online, or through social media. However, if they are to be effective, such interventions cannot simply replicate the content that students already receive in the form of numerous college brochures. As noted above, the two most obvious deficiencies of these brochures are (i) that they are generic rather than customized to a student's situation (for instance, his family's finances) and (ii) that they have a boosterism that may make it difficult for students to derive information from them. Taking these points to heart, we test several interventions in follow-up work (Hoxby and Turner 2012) that has the potential to identify *causal* effects of giving low-income students information about their college-going opportunities.

8 Conclusions

We demonstrate that the majority of high-achieving, low-income students do not apply to any selective colleges despite apparently being well-qualified for admission. These students exhibit behavior that is typical of students of their *income* rather than typical of students of their *achievement*. There are, however, high-achieving, low-income students who apply in much the same way as their high-income counterparts. These "achievement-typical" students also enroll and persist in college like their high-income counterparts. We demonstrate that achievement-typical students come disproportionately from the central cities of large urban areas where they are likely to attend selective, magnet, or other high schools with a critical mass of high achievers.

We note that the majority of achievement-typical students are drawn from only fifteen urban areas, in each of which there is at least one and often several selective colleges. We believe that this phenomenon occurs because many colleges are "searching under the lamp-post." That is, many colleges look for low-income students where the college is instead of looking for low-income students where the students are. The students "under the lamp-post" are already applying to and attending selective colleges.

In contrast, high-achieving, low-income students whose behavior is typical for their income ("income-typical" students) attend schools and live in neighborhoods that lack others who have attended or could attend selective colleges. They are insufficiently geographically concentrated to be reached, cost-effectively, by popular methods of informing students about their college opportunities: visits by admissions staff to high schools, campus visits by students, after school college access programs, contact with teachers who attended selective colleges, and the like.

We speculate that admissions staff believe that the supply of high-achieving, low-income students to selective colleges is inelastic for two reasons. First, many of the students are not on the "radar screen" because they do not apply. Second, the staff spend much of their time informing students who attend high schools that are already "tapped out," so that their efforts offset one another's. That is, their efforts move a well-defined set of students among colleges but do not expand the pool of high-achieving, low-income students who apply to college. Our results suggest that interventions are more likely to affect low-income students' college-going behavior if they do not depend, for their efficacy, on the students being concentrated in a limited number of schools or small geographic areas.

9 References

Avery, C., M. Glickman, C. Hoxby, and A. Metrick. forthcoming. "A Revealed Preference Ranking of American Colleges and Universities," *Quarterly Journal of Economics*.

Avery, C., C. Hoxby, C. Jackson, K. Burek, G. Pope, and M. Raman, "Cost Should Be No Barrier: An Evaluation of the First Year of Harvard's Financial Aid Initiative," NBER WP 12029.

Avery, C. and C. Hoxby, "Do and Should Financial Aid Packages Affect Students' College Choices?" in ed C. Hoxby, *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*. University of Chicago Press: 2004.

Avery, C. and S. Turner. 2009. "Aid and Application Awareness." Unpublished manuscript.

Bowen, W., M. Chingos, M. McPherson. 2009. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton: Princeton University Press.

Card, D. and A.B. Krueger. 2005. "Would the Elimination of Affirmative Action Affect Highly

Qualified Minority Applicants?: Evidence from California and Texas." Industrial and Labor Relations Review, 58(3): 416-434.

Dillon, E. and J. Smith. 2012. "Determinants of Mismatch Between Student Ability and College Quality." University of Michigan unpublished manuscript.

Finn, C. and J. Hockett. 2012. *Shiny Needles in the Education Haystack: America's Academically-selective Public High Schools.* Stanford, CA: Hoover Institution Press.

Freeberg, N. 1988. "Accuracy of Student Reported Information." College Board Report Number 88-5. http://professionals.collegeboard.com/profdownload/pdf/RR%2088-5.PDF

Hill, C., G. Winston, and S. Boyd. 2005. "Affordability: Family Incomes and Net Prices at Highly Selective Private Colleges and Universities." *Journal of Human Resources*, 40(4):769-790.

Hoxby, C. 2010. "The Changing Selectivity of American Colleges and Universities: Its Implications for Students, Resources, and Tuition." *Journal of Economic Perspectives*.

Hoxby, C. and S. Turner. 2012. "Expanding College Opportunities for Low-Income, High-Achieving Students," NBER Conference paper.

Meer, J. and H. Rosen. 2012. "Does Generosity Beget Generosity? Alumni Giving and Undergraduate Financial Aid," NBER Working Paper No. w17861.

Pallais, A. 2009. "Small Differences that Matter: Mistakes in Applying to College." Revise and resubmit, *Journal of Labor Economics*. http://econ-www.mit.edu/files/4030.

Figure 1

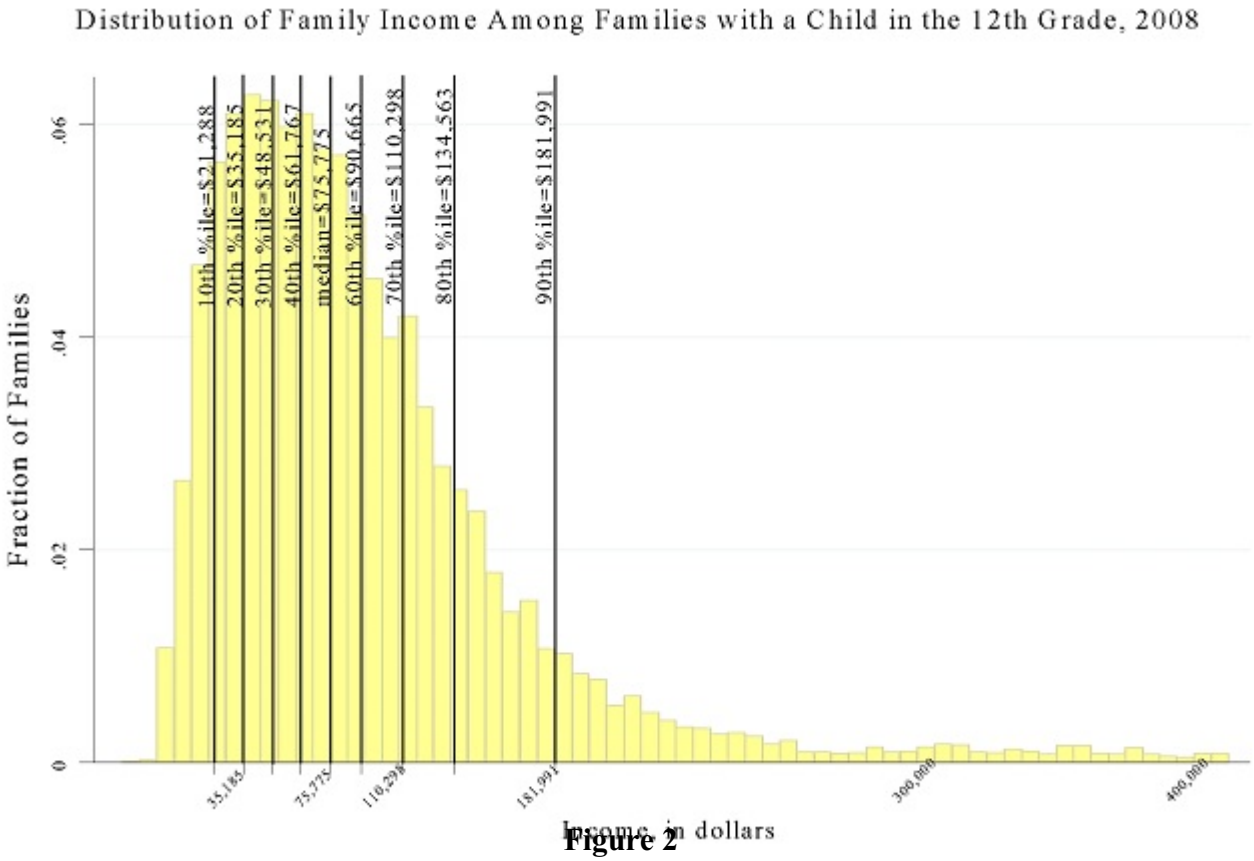

Figure 2

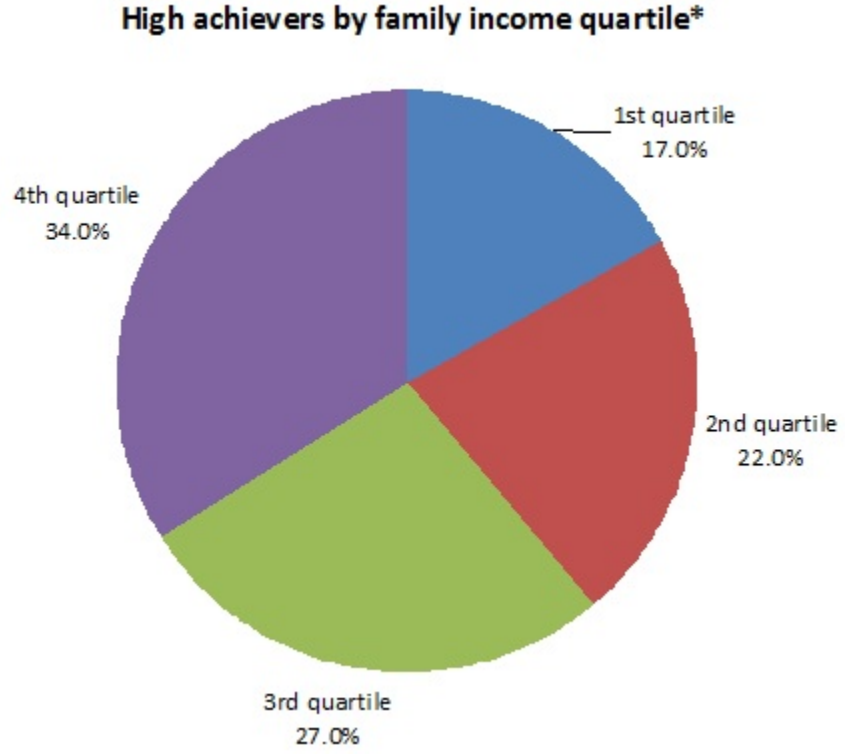

Figure 3

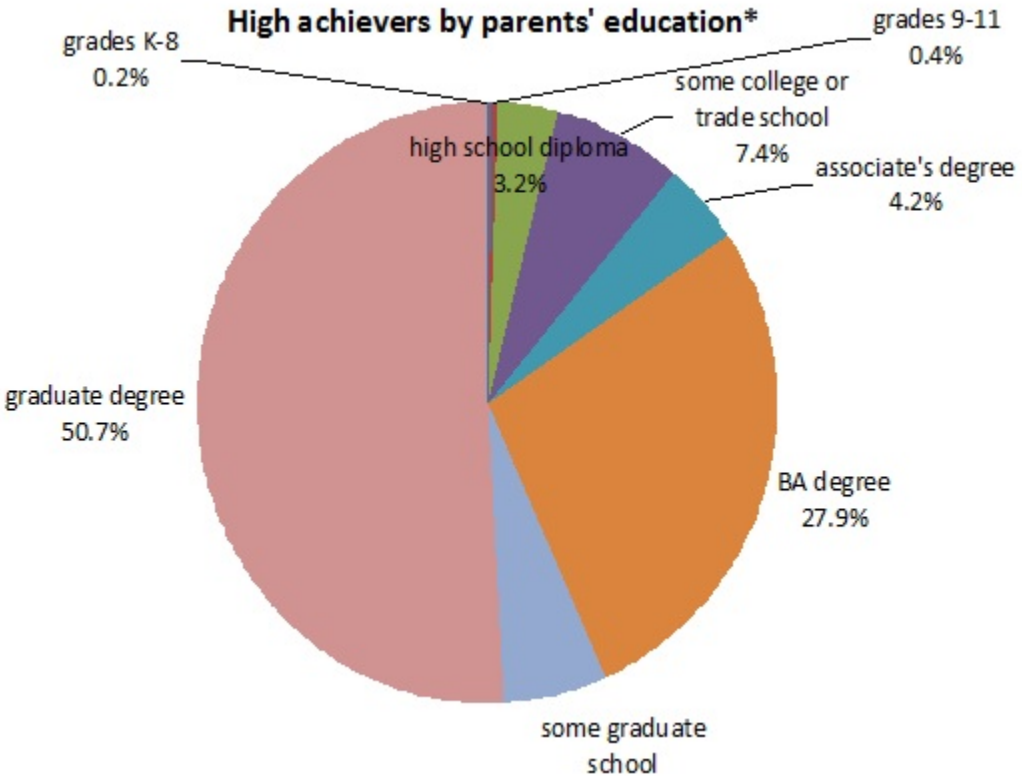

Figure 4

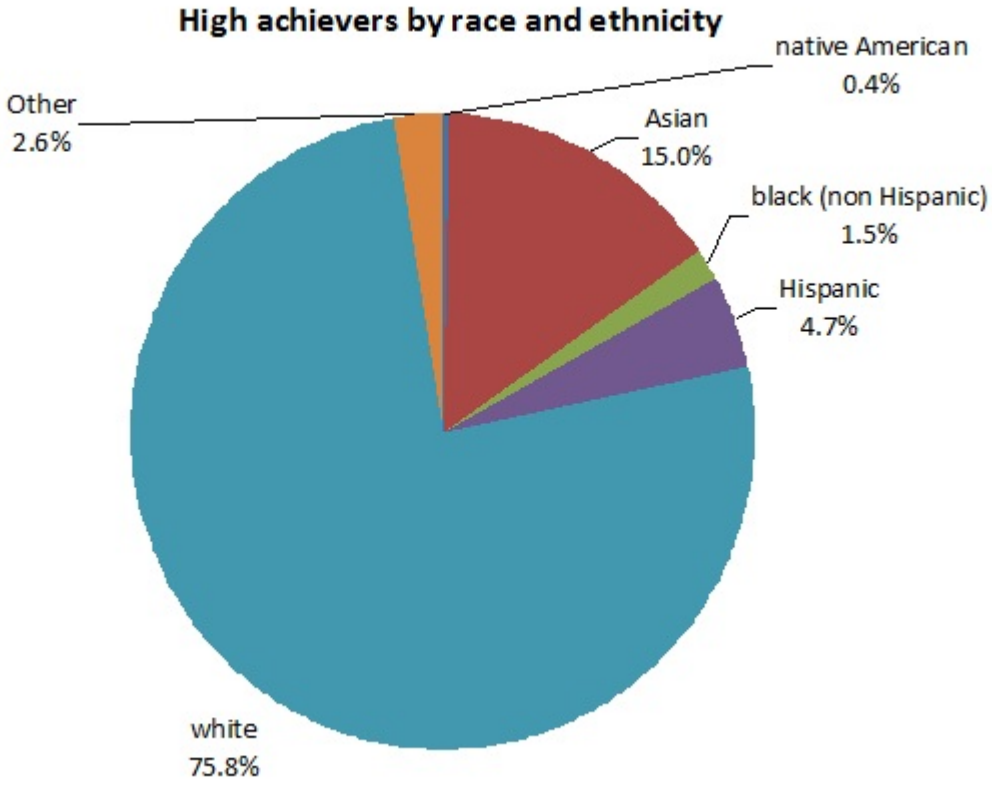

Figure 5

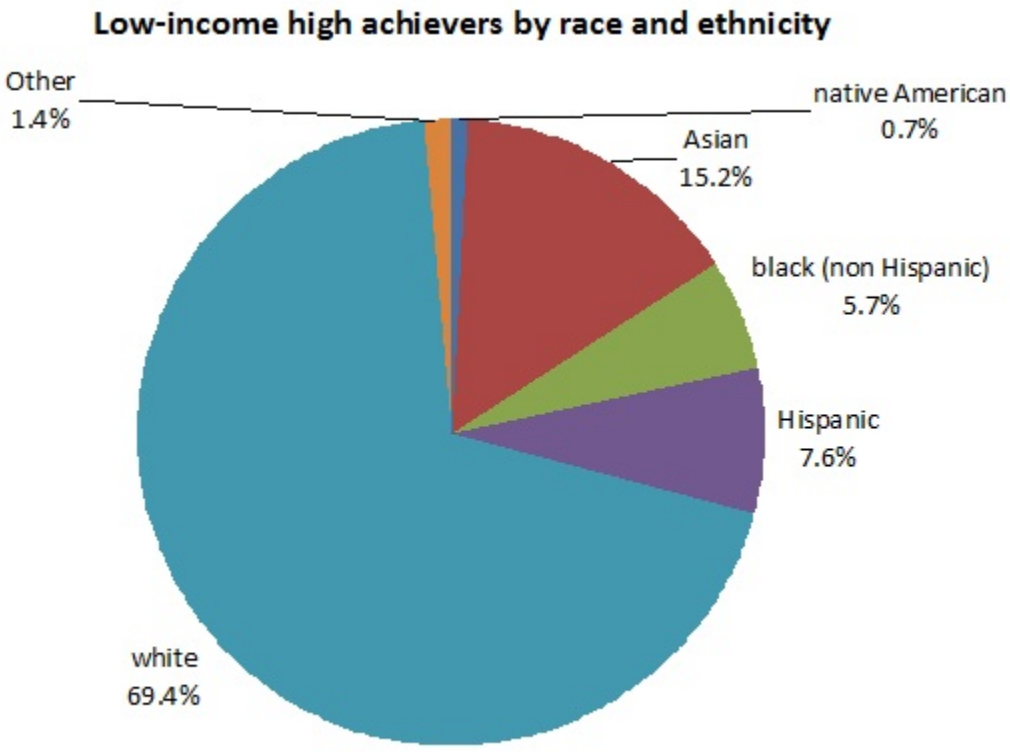

Figure 6

Number of High Achievers by County
darker = a greater number of high achievers

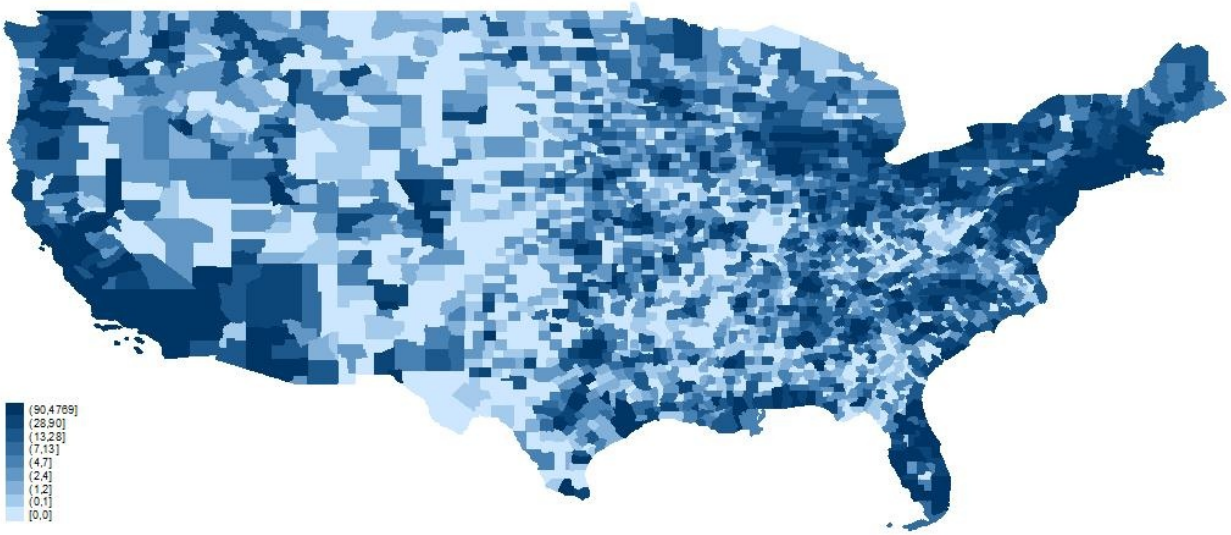

Figure 7
Number of High Achievers per 17-year-old
darker = greater number of high achievers per 17-year-old

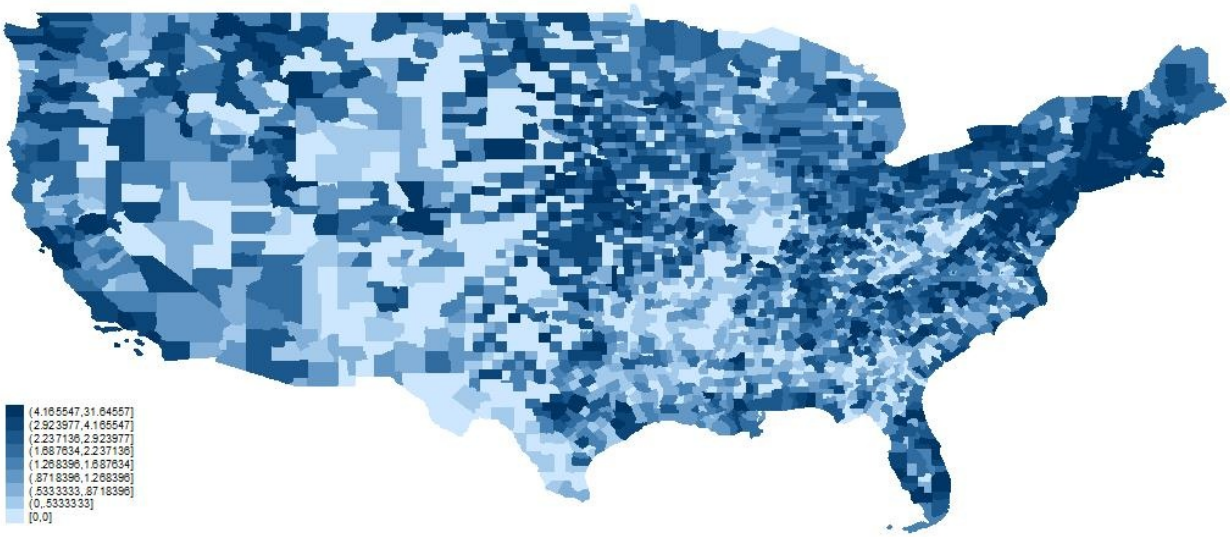

Figure 8
High Income Students' Portfolios of College Applications
(1 student = weight of 1)

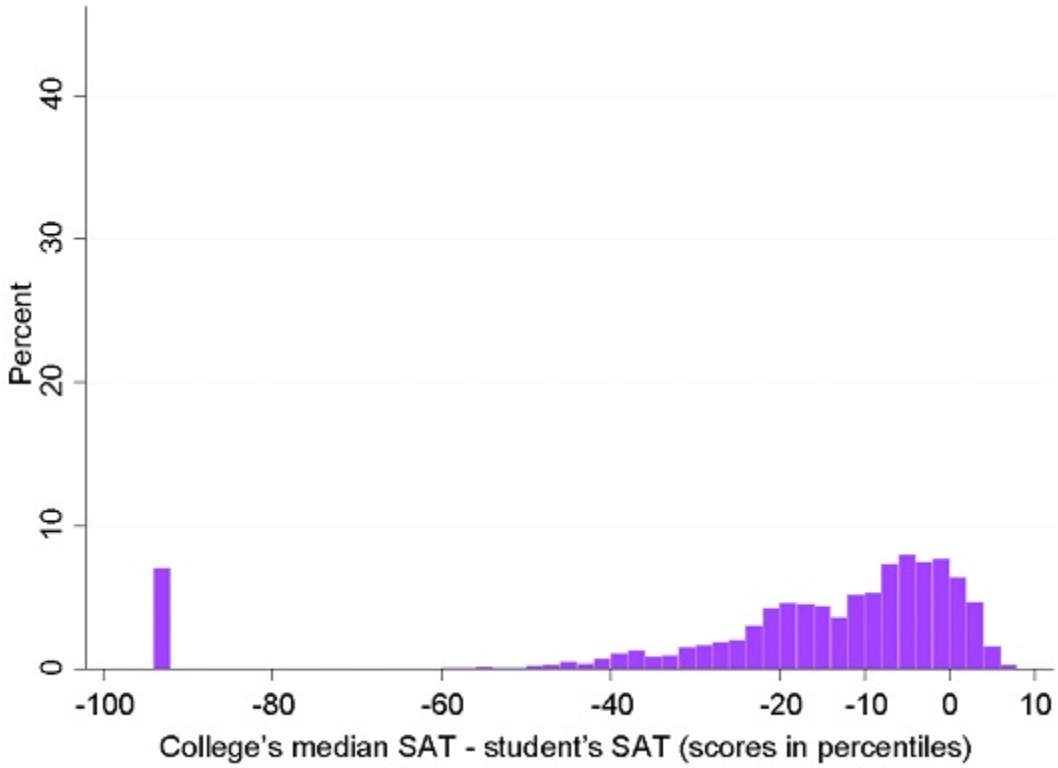

Figure 10
Low Income Students' Portfolios of College Applications
(1 student = weight of 1)

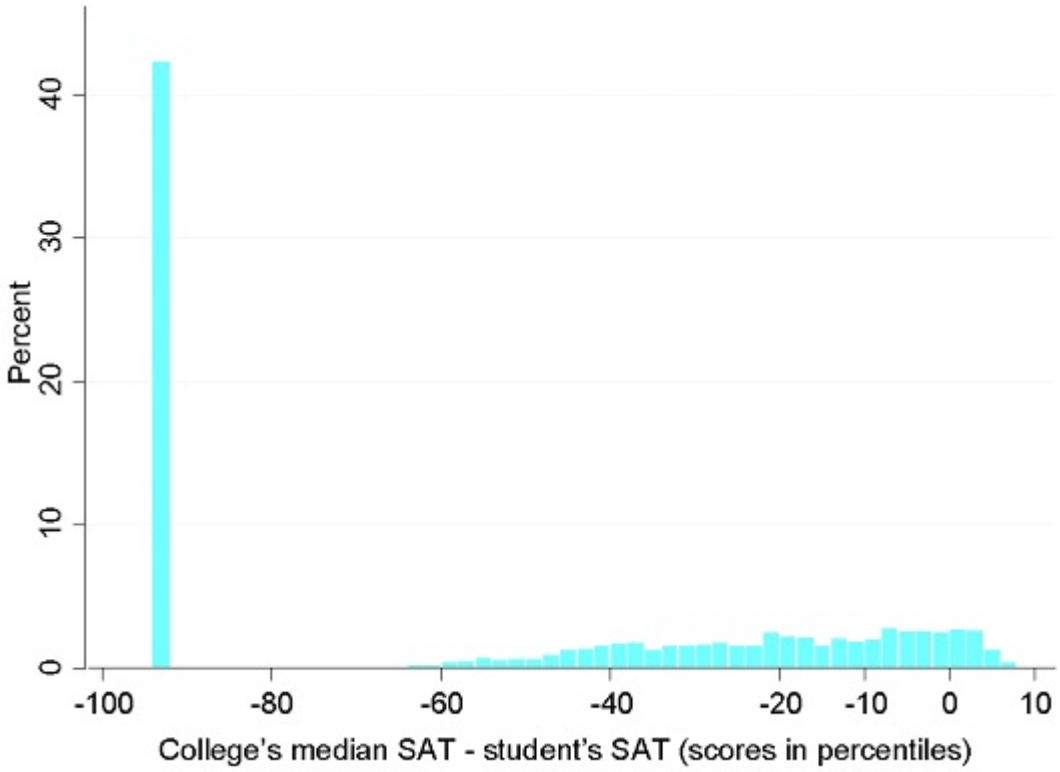

Figure 9
Low, Middle, and High Income Students' Portfolios of College Applications
Excluding Applications to Non-Selective Institutions (1 student = weight of 1)
blue = low income, brown = middle income, purple = high income

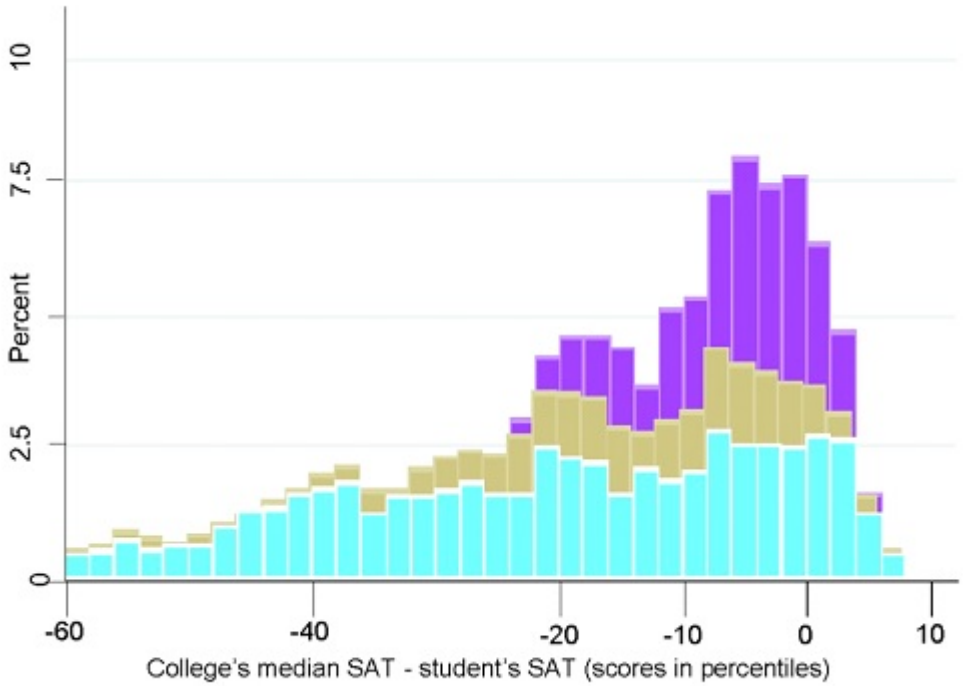

Table 1
College Costs and Resources by Selectivity

Selectivity (Barron's)	Out-of-Pocket Cost for a Student at the 20th Percentile of Family Income (includes room and board)	Comprehensive Cost (includes room and board)	Instructional Expenditure per Student
most competitive	6,754	45,540	27,001
highly competitive plus	13,755	38,603	13,732
highly competitive	17,437	35,811	12,163
very competitive plus	15,977	31,591	9,605
very competitive	23,813	29,173	8,300
competitive plus	23,552	27,436	6,970
competitive	19,400	24,166	6,542
less competitive	26,335	21,262	5,359
some or no selection, 4-year	18,981	16,638	5,119
private 2-year	14,852	17,822	6,796
public 2-year	7,573	10,543	4,991
for-profit 2-year	18,486	21,456	3,257

Notes: The sources are colleges' net cost calculators for the out-of-pocket cost column and IPEDS for the remaining columns. The net cost data were gathered for the 2009-10 school year by the authors, for the institutions at the very competitive and more selective levels. For the institutions of lower selectivity, net cost estimates are based on the institution's published net cost calculator for the year closest to 2009-10--never later than 2011-12. Net costs are then reduced to approximate 2009-10 levels using the institution's own room and board and tuition net of aid numbers from IPEDS, for the relevant years.

Table 2
College Assessment Results of High Achievers, by Family Income

Income Quartile	Average SAT/ACT Percentile among High Achievers
1st quartile (lowest income)	94.1
2nd quartile	94.3
3rd quartile	94.8
4th quartile (highest income)	95.7

Notes: A "high achiever" is student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. The source is authors' calculations based on the combined dataset (ACT, College Board, IPEDS, and other sources) described in the text.

Table 3
Factors Associated with Applying to College
Results of a Conditional Logit Estimation (expressed in odds ratios)

	Low-Income High Achievers	High-Income High Achievers
college is a close match	1.015	76.214***
college is a safety school	3.009***	14.895***
institution is nonselective	0.748***	1.6e-9***
tuition in thousands (sticker price, in- or out-of-state as relevant for individual)	0.865***	1.176***
average tuition discount in percent	1.091**	0.925**
could live at family home (<10 miles)	4.942***	0.810***
could go home often (<120 miles)	1.556***	1.185***
distance in miles	0.996	0.998
$(\text{distance}/1000)^2$	1.056**	1.283***
college is in-state	2.595***	1.206***
college is private	0.838***	1.002
institution is for-profit	0.834***	0.012***
highest degree offered by institution is 2-year	0.925**	0.009***
institution is a university	0.997	0.567***
institution is a liberal arts college	0.717***	0.973*

Notes: The table presents the results of conditional logit estimations in which the dependent variable is an indicator for a high achieving student's having applied to a postsecondary institution. The results are expressed in odds ratios. A "high achiever" is student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. The data source is the combined dataset (ACT, College Board, IPEDS, and other sources) described in the text. Asterisks indicate that the odds ratio is statistically significantly different than 1 at the 1 percent level (***), 5 percent level (**), or 10 percent level (*).

Table 4
Factors Associated with Applying to College,
Income-typical versus Achievement-typical Students
Results of a Conditional Logit Estimation (expressed in odds ratios)

	Income-typical Students	Achievement-typical Students	High-income Students
college is a close match	7.21e-8***	87.808***	76.214***
college is a safety school	2.142***	19.817***	14.895***
institution is nonselective	0.795***	1.04e-8***	1.6e-9***
tuition in thousands (sticker price, in- or out-of-state as relevant for individual)	0.973***	1.004	1.176***
average tuition discount in percent	1.000	1.020*	0.925**
could live at family home (<10 miles)	5.140***	1.477***	0.810***
could go home often (<120 miles)	1.972***	1.436***	1.185***
distance in miles	0.999	0.999	0.998
$(\text{distance}/1000)^2$	1.042*	1.448***	1.283***
college is in-state	4.891***	7.455***	1.206***
college is private	0.662***	0.296***	1.002
institution is for-profit	0.806***	0.001***	0.012***
highest degree offered by institution is 2-year	0.855***	0.016***	0.009***
institution is a university	0.956**	0.861***	0.567***
institution is a liberal arts college	0.515***	0.167***	0.973*

Notes: The table presents the results of conditional logit estimations in which the dependent variable is an indicator for a high achieving student's having applied to a postsecondary institution. The results are expressed in odds ratios. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low- and high-income students are, respectively, ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, IPEDS) described in the text. Asterisks indicate that the odds ratio is statistically significantly different than 1 at the 1 percent level (***), 5 percent level (**), or 10 percent level (*).

Table 5
Estimates Showing Whether Low- and High-Income Students Have Different Probabilities of Enrolling in a College Conditional on Having Applied to It
(results from conditional logit estimation in which coefficients are allowed to differ for low- and high-income students)

College Selectivity	Share of Colleges where Low- and High-Income Students' Probability of Enrollment Conditional on Application is Statistically Significantly Different at the 5 Percent Level
median student is ≥ 90th percentile	4%
median student is ≥80thand <90th percentile	5%
median student is ≥70thand <80th percentile	4%
median student is ≥60thand <70th percentile	3%
median student is ≥50thand <60th percentile	6%
college is selective but median student is <50th percentile	not identified (see notes)
college is non-selective	not identified (see notes)

Notes: The table summarizes results from a conditional logit estimation in which the dependent variable is an indicator for a high achieving student's having enrolled in postsecondary institution. Each student's choice set is the set of colleges to which he applied. The estimating equation's only independent variables are indicators for each college interacted with an indicator for whether the student is high- or low-income. If the coefficients on the high- and low-income indicators for a college are statistically significantly different, then that college is counted in the shares shown in the right-hand column. Each college's results are placed in a row on the basis of the college's selectivity. The results are not identified for low selectivity and non-selective colleges because there are an insufficient number of high-income students who apply to such colleges. That is, for such colleges, the high-income*college-indicator variables are dropped in the process of estimation. A "high achiever" is student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. The data source is the combined dataset (ACT, College Board, IPEDS, NSC, and other sources) described in the text.

Table 6
Estimates Showing Whether Low- and High-Income Students Have Different Probabilities of Persisting in a College Conditional on Having Enrolled
(results from ordinary least squares estimation in which coefficients are allowed to differ for low- and high-income students)

College Selectivity	Share of Colleges where Low- and High-Income Students' Persistence (measured in share of degree completed) Conditional on Enrollment is Statistically Significantly Different at the 5 Percent Level
median student is ≥ 90th percentile	5%
median student is ≥80thand <90th percentile	4%
median student is ≥70thand <80th percentile	4%
median student is ≥60thand <70th percentile	5%
median student is ≥50thand <60th percentile	4%
college is selective but median student is <50th percentile	not identified (see notes)
college is non-selective	not identified (see notes)

Notes: The table summarizes results from an ordinary least squares estimation in which the dependent variable is a student's share of a baccalaureate degree completed by June 2012. Students who do not enroll in postsecondary institution are not included in the regression. The estimating equation's only independent variables are indicators for each college interacted with an indicator for whether the student is high- or low-income. If the coefficients on the high- and low-income indicators for a college are statistically significantly different, then that college is counted in the shares shown in the right-hand column. Each college's results are placed in a row on the basis of the college's selectivity. The results are not identified for low selectivity and non-selective colleges because there are an insufficient number of high-income students who enroll in such colleges. That is, for such colleges, the high-income*college-indicator variables are dropped in the process of estimation. A "high achiever" is student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. The data source is the combined dataset (ACT, College Board, IPEDS, NSC, and other sources) described in the text.

Table 7
Socio-economic Characteristics of High-Achieving Students
Achievement-typical versus Income-typical Students

	High-income	Achievement-typical	Income-typical
Estimated variables:			
family income	$157,569	$30,475	$32,418
Self-reported variables:			
parents' education (see notes)	18.7 years	16.0 years	16.7 years
white	74.80%	45.10%	79.50%
black	2.10%	5.20%	2.90%
Hispanic	5.60%	12.60%	6.00%
Asian	20.50%	31.80%	7.30%

Notes: The table summarizes the characteristics of high-achieving students' families. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, IPEDS, and other sources) described in the text. Self-reported parental education is highly unreliable because students whose parents probably have low education do not report. See text for further explanation.

Table 8
Characteristics of the Neighborhoods of High-Achieving Students
Achievement-typical versus Income-typical Students

	High-income	Achievement-typical	Income-typical
family income (BG)	$123,684	$32,142	$31,767
adjusted gross income (zipcode)	$121,448	$41,358	$37,652
number with a BA (BG)	863	207	144
% with a BA (BG)	66.70%	22.00%	16.80%
% white (BG)	86.70%	58.20%	77.10%
% black (BG)	2.60%	12.80%	10.10%
% Hispanic (BG)	4.10%	16.90%	8.70%
% Asian (BG)	9.20%	8.50%	2.20%

Notes: The table summarizes the characteristics of high-achieving students' Census Block Groups (BG) or zipcodes. These "neighborhoods" are noted in parentheses. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, IPEDS, Geolytics Census Tract, IRS estimates for 2008, and other sources) described in the text.

Table 9
Characteristics of the Home Locations of High-Achieving Students
Achievement-typical versus Income-typical Students

	High-income	Achievement-typical	Income-typical
Main city in urban area w pop 250k+	17%	26%	8%
Main city in urban area w pop 100-250k	14%	21%	13%
Main city in urban area w pop <100k	48%	18%	9%
Suburb in urban area 250k+	8%	9%	9%
Suburb in urban area 100-250k	0%	2%	2%
Suburb in urban pop<100k	0%	4%	12%
Town, near an urban area	0%	5%	12%
Town, far from an urban area	5%	7%	15%
Rural, near an urban area	6%	4%	10%
Rural, far from urban area	0%	5%	10%

Notes: The table summarizes the characteristics of the home locations of high-achieving students. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, Census, and other sources) described in the text.

Table 10
Characteristics of the High Schools of High-Achieving Students
Achievement-typical versus Income-typical Students

	High-income	Achievement-typical	Income-typical
school cohort size	333	330	241
regular public school	66%	73%	86%
magnet school	4%	11%	0%
independent private school	16%	7%	3%
Catholic or other religious school	15%	9%	11%
per pupil spending (public schools)	$15,558	$12,975	$10,701
pupil-teacher ratio (all schools)	16.8	18.3	17.2
pupil-counselor ratio (public schools)	307	328	305

Notes: The table summarizes the characteristics of the high schools of high-achieving students. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, Common Core of Data, and other sources) described in the text.

Table 11
College-Related Characteristics of the High Schools of High-Achieving Students
Achievement-typical versus Income-typical Students

	High-income	Achievement-typical	Income-typical
%teachers graduated from close match college	8.90%	2.90%	1.10%
%teachers graduated from safety college	14.40%	7.50%	5.00%
# applicants to top 10 colleges in each cohort (average of last 10 cohorts)	12.9	7.6	1.6
# admits to top 10 colleges in each cohort (average of last 10 cohorts)	12.3	7.4	1.5
# who enrolled at top 10 colleges each cohort (average of last 10 cohorts)	12.3	7.4	1.5
%high school cohort who are high achievers	17.10%	11.20%	3.80%
radius to gather 20 high achievers	2.6 miles	7.7 miles	19.3 miles
radius to gather 50 high achievers	4.1 miles	12.2 miles	37.3 miles

Notes: The table summarizes college-related characteristics of the high schools of high-achieving students. A "high achiever" is a student with ACT or SAT scores at or above the 90th percentile and a high school grade point average of A- or above. Low-income and high-income students are defined, respectively, as ones from the bottom and top quartile of the family income distribution. Achievement-typical (income-typical) students are low-income, high-achieving students whose application behavior does (does not) resemble that of high-income, high-achieving students. The data source is the combined dataset (ACT, College Board, Common Core of Data, Schools and Staffing Survey, and other sources) described in the text.