Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data

Stacy Dale
Mathematica Policy Research

Alan Krueger
Princeton University

October 19, 2012

Abstract. We estimate the labor market effect of attending a highly selective college, using the College and Beyond (C&B) Survey linked to Detailed Earnings Records from the Social Security Administration (SSA). This paper extends earlier work by Dale and Krueger (2002) that examined the relationship between the college that students attended in 1976 and the earnings they self-reported in 1995 on the C&B follow-up survey. In this analysis, we use administrative earnings data to estimate the effects of several college characteristics (that are commonly used as proxies for college quality) for a more recent cohort: students who entered college in 1989. We also estimate these effects for the 1976 cohort, but over a longer time horizon (from 1983 through 2007).

The effects of college characteristics on earnings are sizeable for both cohorts in regression models that control for variables commonly observed by researchers, such as student SAT scores. However, when we partially adjust for unobserved student ability by controlling for the average SAT score of the colleges that students applied to, our estimates of the effects of college characteristics fall substantially and are generally indistinguishable from zero. There were notable exceptions for certain subgroups. For black and Hispanic students and for students who come from families with less parental education, the effects of college characteristics are large, even in models that adjust for unobserved student characteristics. Finally, contrary to expectations, our results do not suggest the effects of college characteristics increased for students who entered college more recently because estimates are similar for the 1976 and the 1989 cohort.

*We thank the Mellon Foundation for financial support; Ed Freeland for obtaining permission of participating colleges; Matthew Jacobus and Licia Gaber Baylis for skilled computer programming support; Mike Risha of SSA for merging the C&B data with SSA data and for tirelessly running our programs; Lawrence Katz, Sarah Turner, and Mark Dynarski for helpful comments on an earlier draft; and Mark Long for providing data on college SAT scores.

I. Introduction

Students who attend higher quality colleges earn more on average than those who attend colleges of lesser quality. However, it is unclear why this differential occurs. Do students who attend more selective schools learn skills that make them more productive workers, as would be suggested by human capital theory? Or, consistent with signaling models, do higher ability students—who are likely to become more productive workers—attend more selective colleges?

Understanding why students who attend higher quality colleges have greater earnings is crucial for parents deciding where to send their children to college, for colleges selecting students, and for policymakers deciding whether to invest additional resources in higher quality institutions. However, obtaining unbiased estimates of the effects of college characteristics is difficult because of unobserved characteristics that affect both a student's attendance at a highly selective college and his or her later earnings. In particular, the same characteristics (such as ambition) that lead students to apply to highly selective colleges may also be rewarded in the labor market. Likewise, the attributes that admissions officers are looking for when selecting students for college may be similar to the attributes that employers are seeking when hiring and promoting workers.

A wide literature exists on the labor market effects of college characteristics, as summarized in Hoxby (2009) and Hershbein (2011). Many papers have used regression models to control for observed student characteristics, such as high school grades, standardized test scores, and parental background (see, for example, Monks 2000; Brewer and Ehrenberg 1996; Black and Smith 2004), and generally find that attending a higher quality college is associated with higher earnings. However, studies that attempt to adjust for unobserved student quality have reported mixed findings. Dale and Krueger (2002) find that the effect of college characteristics falls substantially after implementing their selection-correction, which partially adjusts for unobserved

student quality by controlling for the average SAT score of the colleges that students apply to and are accepted or rejected by. Hoekstra (2009) uses a regression discontinuity design that compares the earnings of students who were just above the admissions cutoff for a state university to those that were just below it; he finds that attending the flagship state university results in 20 percent higher earnings 5 to 10 years after graduation for white men, but he does not find an effect on earnings for white women. Using an instrumental variables strategy, Long (2008) estimated the local average treatment effect of college quality and did not find a consistent relationship between college characteristics and earnings. Lindahl and Regner (2005) use sibling data to illustrate that the effect of college quality might be overstated if family characteristics are not fully adjusted for because cross-sectional estimates are twice as large as within-family estimates. It is important to note that most of the above literature has used a single college characteristic, such as school average SAT score, expenditures per student, the Barron's index, or whether the student attended a flagship state university, as a proxy for college quality. However, Black and Smith (2006) show that the estimates of the effects of school quality are attenuated when a single measure is used; the effects of composite measures are much higher.

One recent study (Hershbein 2011) has tried to distinguish between human capital models and signaling models by empirically testing the relationship between grade point average (GPA), college selectivity, and wages over time. He finds that the return to GPA is smaller at more selective schools than at less selective schools, which is consistent with signaling models. (The marginal benefit of information about GPA is lower at more selective schools because attending a highly selective college already sends a signal about student ability).

Finally, some papers have examined the returns to college quality over time, both within and across cohorts. These studies have generally found that when later recent are cohorts are

compared to earlier cohorts, the premium to attending college has increased (Brewer et al. 1999; Bound and Johnson 1992; Long 2009; Grogger and Eide 1995; Katz and Murphy 1992). However, Black, Daniel and Smith (2005) show that the effects of college quality for a single cohort—the 1979 cohort of the National Longitudinal Survey of Youth (NLSY)—remain stable over an 11-year horizon.

Little research has examined the effects of college characteristics for recent cohorts. This is a notable gap in the literature; one might expect that it would be more important for students who entered college recently to distinguish themselves by attending more selective colleges because the percentage of students enrolling in college has increased.[1] Those studies that do use recent cohorts tend to model earnings early in the career. For example, Long (2008, 2009) used a relatively recent cohort (the 1992 cohort of the National Education Longitudinal Study [NELS]), but he was only able to examine the earnings of students relatively early in their careers when they were only 26 years old.

In this paper, we examine whether the college that students attend (within a set of somewhat selective to highly selective colleges) affects their later earnings. This paper replicates earlier work that examined the relationship between the college that students attended in 1976 and the earnings they reported in 1995 in the College and Beyond (C&B) follow-up survey (Dale and Krueger 2002) and extends this work in important respects. First, we estimate the effects of several college characteristics that are commonly used as proxies for college quality (college average SAT score, the Barron's index, and net tuition) for a recent cohort of students—those who entered college in 1989. By linking the C&B data to administrative records from the Social Security

[1] For example, the percentage of 18- to 24-year-olds enrolling in college increased from 26 percent in 1975 to 32 percent in 1990 (Fox, Connolly, and Snyder 2005).

Administration (SSA), we are able to follow this cohort for 18 years after the students entered (and 14 years after they likely would have graduated from) college. Second, we estimate the return to college characteristics for the 1976 cohort over a long time horizon, from 1983 to 2007. Because we use administrative earnings records from tax data, our earnings measure is presumably more reliable than much of the prior literature, which is generally based on self-reported earnings. The use of administrative earnings data allows us to follow a recent cohort of students over a longer period of time than is possible in many of the longitudinal databases that are typically used to study the returns to college characteristics. For example, the NELS, High School and Beyond, and the National Longitudinal Study of the High School Class of 1972 (NLS-72) only follow students for 6 to 10 years after students would have likely graduated from college; although the NLSY follows students for a longer period of time, students from the relatively recent cohort (who were age 12 to 16 in 1997) are now too early in their post-collegiate careers to generate meaningful estimates of the labor market effects of college characteristics.

As in the rest of the literature, we find that the effect of each college characteristic is sizeable for both cohorts in cross-sectional least squares regression models that control for variables commonly observed by researchers (such as student characteristics and SAT scores). However, when we adjust for a proxy for unobserved student characteristics—namely, by controlling for the average SAT score of the colleges that students applied to—our estimates for the effects of college characteristics fall substantially and are generally indistinguishable from zero for both the 1976 and 1989 cohort of students. Notable exceptions are for racial and ethnic minorities (black and Hispanic students) and for students whose parents have relatively little education; for these subgroups, our estimates remain large, even in models that adjust for unobserved student characteristics. One possible explanation for this pattern of results is that

highly selective colleges provide access to networks for minority students and for students from disadvantaged family backgrounds that are otherwise not available to them. Finally, contrary to expectations, our estimates do not suggest that the effects of college characteristics (within the set of C&B schools) increased for students who entered college more recently because estimates for the 1976 and 1989 cohort are similar when we compare the estimates for each cohort at a similar stage relative to college entry (approximately 18 to 19 years after the students entered college).

II. Methods

The college application process involves a series of choices. First, students choose where to apply to college; then, colleges decide which students to admit. Finally, students choose which college to attend from among the set of schools to which they were admitted. The difficulty with estimating the labor market return to college quality is that not all of the characteristics that lead students to apply to and attend selective colleges are observed by researchers, and unobserved student characteristics are likely to be positively correlated with both school quality and earnings.

We assume the equation relating earnings to the students' attributes is

(1) $\ln W_i = \beta_0 + \beta_1 Q_i + \beta_2 X_{1i} + \beta_3 X_{2i} + \varepsilon_i,$

where Q is a measure of the selectivity of the college student i attended, X_1 and X_2 are two sets of characteristics that affect earnings, and ε_i is an idiosyncratic error term that is uncorrelated with the other explanatory variables (1). X_1 includes variables that are observable to researchers, such as grades and SAT scores, whereas X_2 includes variables that are not observable to researchers, such as student motivation and creativity (that are at least partly revealed to admissions officers through detailed transcript information, essays, interviews, and recommendations). Both X_1 and X_2 affect the set of colleges that students apply to, whether they are admitted, and possibly which school they

attend. The parameter β_1 represents the gross monetary payoff to attending a more selective college.

Early literature on the returns to school quality was generally based on a wage equation that omitted X_2:

(2) $\ln W_i = \beta'_0 + \beta'_1 Q_i + \beta'_2 X_{1i} + u_i$.

Q_i is typically measured by the average SAT score of the school where the student attended college. Even if students randomly select the college they attend from the set of colleges that admitted them, estimation of (2) will yield biased and inconsistent parameter estimates of β_1 and β_2. If students choose their school randomly from their set of options, the payoff to attending a selective school will be biased upward because students with higher values of the omitted variable, X_2, are more likely to be admitted to and therefore attend highly selective schools. Because the labor market rewards X_2, and Q and X_2 are positively correlated, the coefficient on school quality will be biased upward.

To address the selection problem, we use one of the selection-adjusted models—referred to as the "self-revelation model"—in Dale and Krueger (2002). This model assumes that students signal their potential ability, motivation, and ambition by the choice of schools they apply to. If students with greater unobserved earnings potential are more likely to apply to more selective colleges, the error term in equation (2) could be modeled as a function of the average SAT score (denoted AVG) of the schools to which the student applied: $u_i = t_0 + t_1 AVG_i + v_i$. If v_i is uncorrelated with the SAT score of the school the student attended, one can solve the selection problem by including AVG in the wage equation. This approach is called the self-revelation model because individuals reveal their unobserved ability by their college application behavior. This model also includes dummy variables indicating the number of schools the students applied to (in addition to the average SAT score of the

schools) because the number of applications a student submits may also reveal unobserved student traits such as ambition and persistence.

Dale and Krueger (2002) also estimated a matched applicant model that included an unrestricted set of dummy variables indicating groups of students who received the same admissions decisions (i.e., the same combination of acceptances and rejections) from the same set of colleges. The self-revelation model is a special case of the matched applicant model. The matched applicant model and self-revelation model yielded coefficients that were similar in size, but the self-revelation model yielded smaller standard errors. Because of the smaller sample size in the present analysis, we therefore focus on the self-revelation model.

As discussed in more detail Dale and Krueger (2002), a critical assumption of the self-revelation model is that students' enrollment decisions are uncorrelated with the error term of equation (2) and X_2. Our selection correction provides an unbiased estimate of β_1 if students' school enrollment decisions are a function of X_1 or any variable outside the model. However, it is possible that student matriculation decisions are correlated with X_2. For example, past studies have found that students are more likely to matriculate to schools that provide them with more generous financial aid packages (see, e.g., van der Klaauw 1997). If more selective colleges provide more merit aid, the estimated effect of attending an elite college will be biased upward because relatively more students with higher values of X_2 will matriculate at elite colleges, even conditional on the outcomes of the applications to other colleges. If this is the case, our selection-adjusted estimates of the effect of college quality will be biased upward. However, if less selective colleges provide more generous merit aid (leading students with higher values of X_2 to attend less selective schools), the estimate could be biased downward. More generally, our adjusted estimate would be biased upward

(downward) if students with high unobserved earnings potential are more (less) likely to attend the more selective schools from the set of schools that admitted them.

Finally, it is possible that the effect of attending a highly selective school varies across individuals (that is, β_1 could have an i subscript), and students might sort among selective and less selective colleges based on their potential returns at that college, as in the Roy model of occupational choice. In such a model, our estimate of the return to attending a selective school can be biased upward or downward, and it would not be appropriate to interpret an estimate of β_1 as a causal effect for the average student.

III. Data

A. College and Beyond Data

Our study is based on data from the 1976 and 1989 cohorts of the College and Beyond Survey. The C&B data set includes information drawn from the applications and transcripts of 34 colleges and universities (including 4 public universities, 4 historically black colleges and universities [HBCUs], 11 liberal arts college, and 15 private universities). Much of the past research using the C&B data (such as Bowen and Bok 1998 and Dale and Krueger 2002) excluded the 4 HBCUs.[2] In this analysis, we include the 27 schools (listed in Appendix Table A2) that agreed to participate in this follow-up study, which included 3 public universities, 10 liberal arts colleges, 12 private universities, and 2 HBCUs. Our sample represents 81 percent of the students included in the original C&B data set.

The original C&B Survey, conducted by Mathematica Policy Research (Mathematica) in

[2] At the time that Dale and Krueger (2002) was written, the HBCUs were not part of the standard C&B data set that was provided to researchers.

1994–1996, contained questions about earnings, occupation, demographics, education, civic activities and life satisfaction.[3] Mathematica attempted to survey all students in the 1976 cohort from each of the 34 C&B schools, with the exception of the four public universities, where a sample (of 2,000 individuals) was drawn that included all racial and ethnic minorities and athletes along with a random sample of other students. For the 1989 cohort, students from 21 colleges were surveyed (listed in Appendix Table A2). The original 1989 C&B sample included all racial and ethnic minorities and athletes and a random sample of other students. Our regressions are weighted by the inverse of the probability that a student was included in the sample.

Early in the C&B questionnaire respondents were asked, "In rough order of preference, please list the other schools you seriously considered."[4] Respondents were then asked whether they applied to, and were accepted by, each of the schools they listed. Because our analysis relies on individuals' responses to these survey questions, our primary analysis is restricted to survey respondents.[5] Survey response rates were 80 percent for the 1976 cohort and 84 percent for the 1989 cohort.

---

[3] See Bowen and Bok (1998) for a full description of the C&B data set.

[4] Students who responded to the C&B pilot survey were not asked this question and are therefore excluded from our analysis.

[5] We were able to estimate our basic wage equation for the full sample of C&B students (including nonrespondents) and obtained results that were similar to those restricted to survey respondents. For example, if we include all students in the 1976 cohort with non-zero earnings, the coefficient on school SAT score in the 1995 earnings basic regression model was 0.059 with a standard error of 0.021; for the sample of survey respondents with non-zero earnings, the coefficient on school SAT score was 0.061 with a standard error of 0.019 (not shown).

The C&B Survey data were drawn from individuals' college applications (such as their SAT scores) and transcripts (such as grades in college). The C&B data were also merged to the Higher Education Research Institute's (HERI) Freshman Survey.

B. Regression Control Variables

Our basic regression model controls for race, sex, high school GPA, student SAT score as reported on the student's application (generally the highest), predicted parental income, and whether the student was a college athlete; our self-revelation model includes these same variables and also the average SAT score of the schools to which a student applied and the number of applications he or she submitted. Our models generally only include the main effects for each of the control variables, though we test one set of models that interacts parental education with college characteristics.[6] Race, gender, parental education and occupation (used to predict parental income), information on the schools the student applied to, whether the student was an athlete, and student SAT score were drawn from the C&B data. To construct other variables about students' performance in high school and their parents' income, we used data from the HERI freshman survey. Because the HERI survey was not completed by all students in the C&B sample, about half of the sample was missing GPA (see Table 1) and parental income. For parental income, we constructed an index that captures a student's family background information.[7] To do this, we first regressed log parental income on

[6] Because we relied on SSA to run programs for us (and did not have access to SSA data), we used a parsimonious regression specification. In exploratory analyses for Dale and Krueger (2002), we found that the effects of college characteristics were generally not sensitive to the coding of regression control variables.

[7] Analyses conducted using the C&B data for Dale and Krueger (2002) suggested that estimates of the effects of college characteristics were not sensitive to whether the underlying components of

mother's and father's education and occupation for the subset of students with available family income data and then multiplied the coefficients from this regression by the values of the explanatory variables for every student in the sample. When regression control variables for SAT score or high school GPA were missing, we set the variable equal to the mean value for the sample and also included a dummy variable indicating the data were missing.

C. College Characteristics

Each college's average SAT score and Barron's index of college selectivity (as reported in the 1978 and 1992 editions of *Barron's Profiles of American Colleges*) were linked to student's responses to the questions concerning the schools they applied to.[8] Because there were only one or two colleges in some categories of the Barron's index (particularly for the 1989 cohort), we represent the index with a continuous variable that ranges from 2 (Competitive) to 5 (Most Competitive) in our sample (College Division of Barron's Education Series 1978, 1992).

Net tuition for 1970, 1980, and 1990 was intended to capture the average amount students paid to attend a particular college.[9] We calculated this measure by subtracting the average aid awarded to undergraduates from the sticker price tuition, as reported in the 11th, 12th, and 14th editions of *American Universities and Colleges* (American Council on Education 1973, 1983, 1992). The 1976 net tuition was interpolated from the 1970 and 1980 net tuition, assuming an

---

predicted parental income (education and occupation) were included as regression control variables in place of this index.

[8] Files with average SAT scores were provided by HERI (for 1978) and by Mark Long (for 1992).

[9] Although not a direct measure of college quality, one might expect that students and their parents would be willing to pay a higher net tuition for colleges that are most likely to increase the student's future earnings potential.

exponential rate of growth. The correlation between these measures were high: 0.81 between net tuition and school SAT score, 0.91 between the Barron's index and college average SAT score, and 0.86 between the Barron's index and net tuition.

D. Earnings Measures

The Social Security Administration linked C&B data to SSA's Detailed Earnings Records for the period of 1981 through 2007. The earnings measure for this analysis included the total earnings an individual reported to the Internal Revenue Service, including earnings from self-employment and earnings that were deferred to retirement plans (but excluding income from capital gains). SSA ran computer programs written by Mathematica on our behalf so that individual-level earnings data were never viewed by researchers outside SSA. By using Social Security numbers, SSA was able to match more than 95 percent of the student records we provided. We converted annual earnings for each year to 2007 dollars using the Consumer Price Index. The SSA earnings measure used in our primary analysis is not topcoded; however, to compare to the C&B survey, for one analysis, we deliberately topcoded the SSA data to be consistent with the C&B data (as described below).

For some analyses, we use outcome measures that were the median of an individual's log annual earnings in 2007 dollars over five-year intervals (1983 through 1987, 1988 through 1992, 1993 through 1997, 1998 through 2002, and 2003 through 2007). For example, the dependent variable for the period of 1993 to 1997 was the median (for each individual) of his or her log earnings in the five years from 1993 to 1997. By using medians over five-year intervals, we reduce noise in the earnings measure that would result from brief periods of time that the students may have spent out of the labor market or in non-covered employment.

Finally, consistent with most of the literature, the focus of this study is on the earnings of individuals who are employed (and not on whether individuals choose to or are able to work). Because we cannot identify full-time workers or hourly wages in the SSA administrative data, we generally restrict the sample to those earning more than $13,822 (in 2007 dollars) during the year, the equivalent of earning the minimum wage for 2,000 hours at the 1982 federal minimum wage value (in 2007 dollars). For those regressions in which the dependent variable is median earnings over a five-year interval, individuals were included in the sample if their median earnings over the five-year interval exceeded $13,822; individuals were still included in the sample if they earned less than $13,822 in a particular year as long as their median earnings exceeded $13,822. Estimates based on a sample that use this restriction are more precise than those based on a sample of all non-zero earners.[10] Also, as shown in Table 3, estimates based on the sample defined by this restriction are closer to estimates drawn from the sample of full-time workers (according to the C&B survey) than are estimates drawn from a sample of all non-zero earners because using the minimum wage threshold allows us to exclude those who are clearly not working full time.[11]

---

[10] Approximately 10 percent of workers in our sample (that is, those with any earnings) in the 1976 cohort and 8 percent of those in the 1989 cohort had earnings that were between zero and this minimum wage threshold ($13,822).

[11] Most studies on the return to college quality either restrict the sample to full-time workers (for example, Long 2008) or to non-zero earners (for example, Hoekstra 2009). If we estimate our model using levels instead of logs, and include those with no earnings, we obtain qualitatively similar results. For example, for the 1976 cohort, the parameter estimate (and standard error) for college SAT score was $26,575 (7,566) in the basic model and fell to $2,154 (9,884) in the self-revelation model.

Table 4 helps to assess how these sample restrictions may have affected our results. There does appear to be a negative relationship between attending a school with a higher average SAT score and having wages above the minimum wage threshold, as shown in our basic model in the top panel of Table 4; however, this relationship is statistically insignificant in the self-revelation model. Similarly, individuals who attend colleges with higher average SAT scores are less likely to have non-zero earnings (bottom panel, Table 4). These results suggest that the effects of college characteristics would be attenuated, particularly in the basic model, if we had we included those with no earnings or very low earnings in our regressions (consistent with what is shown by comparing columns 9 to 11 in Table 3).[12]

IV. Descriptive Statistics for Schools and Students

A. Characteristics of Colleges and Students in Sample

Although the average SAT score for colleges in the C&B data set ranged from approximately 800 to over 1300, most of the C&B schools were highly selective. The majority of C&B colleges fell into one of the top two Barron's categories (Most Competitive or Highly Competitive; see Appendix Table A1) and had an average student SAT score of greater than 1175. To place this in context, for students in the NLS-72 sample (a nationally representative database), the college average SAT score was 994 and only 5 percent of college students attended a college or university that fell into one of the top two Barron's categories. The high selectivity of the colleges within the C&B database make it particularly well suited for this analysis because the majority of

[12] In sensitivity tests of the basic model for the 1989 cohort, the coefficient and standard error on school SAT score is 0.034 (0.018) when we include all non-zero workers, compared to 0.056 (0.014) when we restrict the sample to those over the minimum wage threshold.

students that attend selective colleges submit multiple applications, which is necessary for our identification strategy. In contrast, many students who attend less selective colleges submit only one application, because many less selective colleges accept all students who apply. For example, according to data from the NLS-72, only 46 percent of students who attended college applied to more than one school.

The regression sample includes students who entered (but did not necessarily graduate from) one of the C&B schools. Because the schools included in the database were highly selective, the students who were in the sample had high academic qualifications. The students in the 1976 cohort had an average SAT scores of 1160 and an average high school grade point average of 3.6 (Table 1). (Note that for ease of interpretation, in our tables and regression analysis, we divide our measures of school SAT score and student SAT score by 100.) Similarly, for the 1989 cohort, the average student SAT score was greater than 1200, and the average GPA was 3.6. The percentage of students that were racial and ethnic minorities was higher for the 1989 cohort (where 8 percent were black and 3 percent were Hispanic) than for the 1976 cohort (where 6 percent of students were black and 1 percent were Hispanic). Finally, earnings for the sample were high: the average of each individual's median earnings over the 2003 to 2007 period was $164,009 for the 1976 cohort. Average annual earnings in 2007 were $183,411 for the 1976 cohort and $139,698 for the 1989 cohort.

B. Application and Matriculation Patterns

Table 2 provides descriptive statistics about the application behavior of the students who entered one of the C&B schools in our study in 1976 or 1989. Nearly two-thirds of the 1976 cohort and 71 percent of the 1989 cohort submitted at least one additional application (in addition to the school they attended). For both cohorts, of those students submitting at least one additional

application, more than half applied to a school with a higher average SAT score than that of the college they attended, and nearly 90 percent of these students were accepted to at least one additional school. Of those accepted to more than one school, about 35 percent were accepted to a school with a higher average SAT score than the one they ended up attending, with about 23 percent being accepted to a school with an average SAT score that was at least 40 points higher than the one they attended. Blacks and Hispanic students were somewhat more likely than students in the full sample to be accepted to at least one additional school and to be accepted to a more selective school than the one they attended (columns 2 and 4).

Although we could not explore whether students' unobserved ability is related to the school they attended, we were able to examine how students' observed characteristics are related to the school they attended. Predicted parental income, student SAT score, and high school grade point average all show a high, positive correlation with the average SAT score of the college attended (see Appendix Table A4). We also examined the relationship between student characteristics and the average SAT score of a school they chose to attend, conditional on the average SAT score of the most selective school to which they applied (Appendix Table A4). For 1976, the coefficient on student SAT score and high school GPA is positive and statistically significant. These results suggest that students in the 1976 cohort with better academic credentials tended to matriculate to more selective schools, controlling for the average SAT score of the most selective school to which they applied. If, among students who apply to similar schools, more ambitious students choose to attend more selective schools, then even our selection-adjusted estimates of the effect of college selectivity for the 1976 cohort will be biased upward. For the 1989 cohort, however, there was not a consistent pattern between student characteristics and students' choice of schools. Although the relationship between the student's SAT score and the

SAT score of the school the student attended was positive and statistically significant, the relationship between high school GPA and the SAT score of the college attended was negative and statistically significant. Also, for the 1989 cohort, the relationship between predicted parental income and the average SAT score of the college attended was positive and statistically significant.

Another factor that would be expected to influence student matriculation decisions is financial aid. By definition, merit aid is related to the school's assessment of the student's potential. If more selective colleges provide more merit aid, the estimated effect of attending an elite college will be biased upward. On the other hand, if more selective colleges offer more need-based aid, and family income is not perfectly captured in our regression model, then it is possible that the relationship between college characteristics and student earnings will be biased downward. The limited financial aid data available (for a subset of students and schools) suggest that receiving financial aid was correlated with attending colleges with higher average SAT scores, though we were unable to systematically distinguish between need-based and merit-based aid.

V. Results

A. Comparison of Earnings Using C&B Survey and SSA Administrative Data

We begin by comparing earnings data drawn from the C&B survey to those drawn from SSA administrative data. The C&B survey asked individuals to report their earnings in categories; we assigned those individuals with earnings greater than $200,000 a topcode of $245,662. (This topcode was set to be equal to the mean log earnings for graduates age 36 to 38 who earned more than $200,000 per year, according to data from the 1990 census (in 1995 dollars).) If we recode the SSA data so that those earning more than $200,000 have this same topcode, the correlation for the

1976 cohort between SSA earnings (in 1995) and C&B earnings during the same year is 0.90.[13] This is similar to estimates of the reliability of self-reported earnings data in Angrist and Krueger (1999).

To compare results from this analysis to the results reported in Dale and Krueger (2002), we first estimated a regression where the log of C&B earnings is the outcome measure, but restricted the sample to students in the merged C&B and SSA sample (that is, they matriculated at one of the C&B schools participating in this study, reported that they were working full time during all of 1995 on the C&B survey, and matched to the SSA data). The coefficient on school SAT score/100 in the basic model using this sample restriction is 0.068 (0.014) (see Table 3, column 3), indicating that attending a school with a 100-point higher SAT score is associated with approximately 7 percent higher earnings later in a student's career. This estimate is similar (though slightly less than) the 0.076 (0.016) estimate for the C&B sample reported in Dale and Krueger (2002) (shown here in column 1).[14] In both samples, the return becomes indistinguishable from zero in the self-revelation model (shown in columns 2 and 4).

Next, we use earnings drawn from the SSA data. In column 5 of Table 3, we use the same sample of full-time workers but use SSA earnings that were topcoded in the same way that earnings in the C&B survey were topcoded. In column 7, we use SSA earnings and use the same sample of full-time workers but do not topcode the data. In column 9, we use the log (median of 1993 earnings through 1997 earnings) in 2007 dollars as our outcome measure and restrict the

[13] This correlation falls to 0.67 if SSA earnings are not topcoded.

[14] The estimates from columns 1 and 2 are based on students from 30 C&B schools (all of the C&B schools except for the HBCUs); the column 3 estimate includes the 27 C&B schools participating in this study.

sample to those with non-zero earnings. In column 11, we restrict the sample to those with annual earnings that were greater than a minimum wage threshold (defined as $13,822 in 2007 dollars). In each model, the estimates for the coefficient on school SAT score drawn from our basic model range from 0.048 to 0.061 and are similar to (but somewhat less than) the estimate using earnings from the C&B survey as the outcome measure.

Columns 6, 8, 10, and 12 show results from the self-revelation model for each of these samples. The effect of school SAT score in each of these selection-adjusted models is negative and indistinguishable from zero.

In summary, for the 1976 cohort, across a variety of sample restrictions and across both sources of earnings data (C&B survey data and SSA administrative data), the effect of school SAT score is large and positive when we do not adjust for unobserved student characteristics. However, in the self-revelation model, when we include the average SAT score of the schools the student applied to as a control variable—which partially adjusts for unobserved student characteristics—the effect falls substantially, becoming indistinguishable from zero.

B. Alternative Selection Controls

We also reestimated the series of models from Dale and Krueger (2002) that use a variety of selection controls in place of the average SAT scores of the schools that the student applied to. For example, in one model, we controlled for the highest SAT score of the schools a student was accepted by but did not attend. In another model, we controlled for the average SAT score of the colleges that rejected the student. Consistent with Dale and Krueger (2002), in each of these models, the return to the school SAT score of the school that the student actually attended was less than the return to the colleges he or she applied to but did not attend. In models that control for the

average SAT score of the colleges that students were accepted at (in addition to the average SAT score of the colleges the student applied to), the estimated return to college characteristics tends to be slightly lower than in models that only control for the colleges that the students applied to. This is likely because students that are accepted to colleges with higher average SAT scores have higher unobserved ability than those that apply but were not accepted. Finally, the effect of school SAT score falls only modestly if the only additional control variables we add to the basic model are the number of applications the student submitted. In this type of model, the coefficient on school SAT score tends to fall from about 0.07 in the basic model to about 0.06 in the selection-adjusted model; thus, a key part of our selection adjustment includes controlling for the average SAT score of the colleges the student applied to.[15]A full set of these results is available upon request.

C. Estimated Effect of College Characteristics Over the Life Cycle for the 1976 Cohort

To assess the return to school characteristics over the course of a student's career for the 1976 cohort, we estimate regressions where the outcome measure was the median of log(annual earnings) for each individual (in 2007 dollars) over a five-year interval (1983 through 1987, 1988 through 1992, 1993 through 1997, 1998 through 2002, and 2003 through 2007). In our basic model with a standard set of regression controls, the return to college SAT score increases over the course of a student's career, from indistinguishable from zero for the earliest period (1983 to 1987, about three to seven years after students likely would have graduated) to more than 7 percent for the period of 2003 to 2007 (23 to 27 years after college graduation; Table 5). However, in our

[15] If we control only for demographic information (race and gender), the coefficient on school SAT score is about .10, but this coefficient falls as each additional control variable (predicted parental income, SAT score, and high school GPA) is added.

self-revelation models, the estimates are not significantly different from zero for any time period. (To save space, we only report parameter estimates for school characteristics in these tables. In Appendix Table A3, we report a full set of parameter estimates for selected models.)

We also estimated regressions separately by gender. In the basic model, the return to college SAT score for men was about 5 percent in 1988 to 1992 and increased over time, reaching a high of nearly 9 percent for the period of 1998 through 2002. For women, the effect of school SAT score was consistently less than the effect for men, ranging from 3 percent (in 1988 to 1992) to nearly 6 percent (in 2003 to 2007). The smaller effect for women does not appear to be solely because we cannot identify which women were working full time in SSA's administrative data; the effect of school SAT score on earnings for women (5 percent) was also smaller than the effect for men (7 percent) in the C&B survey when we limited the sample to those who reported working full time. For both men and women, the coefficient was zero (and sometimes even negative) in the self-revelation model.[16] To increase sample size and improve the precision of our estimates, we

[16] This lower return to college selectivity for women is consistent with other literature. Results from Hoekstra (2009), Black and Smith (2004), and Long (2008) all suggest that the effect of college selectivity on earnings is lower for women than for men. Also, although the coefficients for school SAT in the self-revelation model were negative and significant for women in some years (1993 to 1997 and 1998 to 2002), the pattern of results across all of the models we estimated (which included, for example, different measures of college quality and different minimum wage thresholds) did not suggest that the return for women was significantly less than zero. For example, the coefficients for the Barron's index for women for women was 0.051 (0.011) in the basic model and 0.010 (0.022) in the self-revelation model in 1993 to 1997; similarly, in 1998 through 1992, the coefficient was 0.050 (0.008) in the basic model and -0.004 (0.027) in the

focus on results based on the pooled sample of men and women together throughout the rest of the paper.

We estimated these same regressions for two other college characteristics, the Barron's index and the log of net tuition. The results are summarized in Table 6. In our basic model, the estimated impact of these school characteristics increased over the course of the student's career, with the coefficient on log tuition reaching a high of 0.14 and the Barron's index reaching 0.08 in the last five-year interval (last set of rows, Table 6).[17] However, in the self-revelation model, the estimates fall substantially and are statistically insignificant at the 0.10 level.[18]

---

self-revelation model.

[17] In exploratory analyses with the C&B data, we combined the measures of college quality with the instrumental variables approach suggested by Black and Smith (2006); specifically, we first predicted school SAT score from net tuition and the Barron's index and then estimated the effect of predicted school SAT score on earnings. The coefficient on predicted school SAT score was high: 0.126 with a standard error of 0.011 (compared to an estimate of 0.074 with a standard error of 0.016 if we use actual SAT score). However, the estimates fell substantially in our selection-adjusted models to an estimate of 0.044 (with a standard error of 0.012) when we control for the quality of schools the student applied to and to -0.028 with a standard error of 0.030 if we control for the quality of the colleges the students were accepted by.

[18] We probed the sensitivity of the estimates by including dummy variables for categories (such as Most Competitive) for the gradations of the Barron's index. The estimates for the most selective categories were sizeable and significant compared with the base group of the least selective

These results are partly a contrast to Dale and Krueger (2002), in that the earlier analysis of self-reported earnings data showed a statistically significant relationship between earnings and the log of net tuition in the self-revelation model because the coefficient on net tuition was of 0.058 (0.018). To attempt to reconcile these results with Dale and Krueger (2002), we reestimated the effect of net tuition on self-reported earnings for full-time workers from the C&B survey in 1995 using the subset of students from the schools participating in this study, and found that the coefficient (adjusted for clustering) on log (net tuition) from the self-revelation model was somewhat smaller: 0.041 (0.038), and not statistically significant. When we estimated the same regression for the same sample but used SSA's administrative earnings data in 1995 (instead of self-reported earnings data from the C&B survey), the coefficient (standard error) on net tuition was even smaller: 0.033 (0.046). Moreover, over the full study period (1983 to 2007) the coefficient on net tuition was generally between 0 and 0.02 (and never greater than 0.033) in the self-revelation model based on earnings drawn from SSA administrative data as the outcome measure. Thus, the effect of net tuition based on the single year of self-reported earnings reported in Dale and Krueger (2002) appears to been atypically high relative to the series of estimates we were able to generate using SSA's administrative data, though the large standard errors make it difficult to draw inferences.

C. Estimated Effects of College Characteristics for the 1989 Cohort

Unlike the 1976 cohort, where we have data for most of the student's career, we only have a limited number of post-college years for the 1989 cohort. As shown for the 1976 cohort, there is

---

schools in the basic model but were small and statistically insignificant in the self-revelation model. See Appendix Table A5 for these results.

no return to college characteristics in the early part of a student's career, possibly because many graduates from highly selective colleges attend graduate school and thus forego work experience early in their careers. Therefore, for the 1989 cohort, we focus on the most recent year with earnings data available, 2007, when the students were on average 35 years old. Although the 1989 cohort is too young for us to assess changes in the return to school selectivity over the student's career, results for this cohort do allow us to assess whether estimates for the return to school selectivity are similar across cohorts at one point in the lifecycle.

In 2007, the coefficient for school SAT score/100 was 0.056 with a standard error of 0.014 (or 0.031 if we adjust for clustering among students who attended the same schools) in the basic model (Table 7). Consistent with the results for the 1976 cohort, the coefficient was indistinguishable from zero (-0.008 with a standard error of 0.019) in the self-revelation model. The results for each gender are also similar to those of the 1976 cohort: the coefficient for women (0.032) was lower than the coefficient for men (0.067) in the basic model, and in the self-revelation model, estimates for both men and women are indistinguishable from zero (not shown). The results for the Barron's index were consistent with the results for school SAT score. Specifically, the return to the Barron's index was nearly 7 percent in the basic model but was close to zero in the self-revelation model. For net tuition, our estimates from both models were negative and had large standard errors.[19]

[19] The negative coefficient for net tuition for the 1989 cohort is at least partly driven by liberal arts colleges with high net tuition. When we added a dummy variable for liberal arts colleges as a regression control variable, the coefficient (and standard error) on net tuition in the basic model was 0.061 (0.038) and -0.035 (0.041) in the self-revelation model. (In contrast, adding a liberal arts

D. Estimated Effect of College Characteristics for Racial and Ethnic Minorities

Because some past studies have found that the return to college selectivity varies by race (Behrman, Rosenzweig, and Taubman 1996; Long 2009; Loury and Garman 1995), we also examined results separately for racial and ethnic minorities. To increase the sample size, we pooled blacks and Hispanics together because both groups often receive preferential treatment in the college admissions process (Bowen and Bok 1998). For the 1976 cohort, the return to each college characteristic increased over the course of the student's career, and the magnitude of the coefficients did not fall substantially in the self-revelation model. The coefficient was most pronounced for the Barron's index, where going to a school in a higher Barron's category translated to 6.7 percent higher earnings (for the period of 2003 to 2007) in the basic model and 6.2 percent higher earnings in the self-revelation model; however, the estimate in the self-revelation model was not statistically significant at the 0.10 level because of large standard errors (not shown, but available upon request).

For the 1989 cohort, parameter estimates for each college characteristic ranged from 6.3 for the Barron's index to 17.3 percent for the log of net tuition (Table 8). These estimates remained large in the self-revelation model, ranging from 4.9 for the Barron's index to 13.8 for the log of net tuition. Although the standard errors are also large, some of the estimates are significantly greater than zero. For example, the coefficient on school SAT score/100 was 0.076 with a standard error of 0.032 (or 0.042 after accounting for clustering of students within schools).

Because the historically black colleges and universities in this sample had lower average SAT scores (and lower Barron's indices and net tuition) than did the rest of the institutions in the

---

dummy did not qualitatively change our findings for the return to college average SAT score.)

C&B database, we investigated whether the large return to school selectivity for minority students (but not for all students) was due to the greater range in school selectivity observed for minority students.[20] Specifically, we reestimated the regressions but excluded the HBCUs from the sample. For the 1976 cohort, the return to the Barron's index fell from 6.2 percent to 1.6 percent (and indistinguishable from zero) when we imposed this sample restriction. However, for the 1989 cohort, the estimates for minority students remained large when we excluded the HBCUs, implying returns of 12 percent for attending a school with 100 point higher SAT score and 14 percent for attending a school in a higher Barron's category, even in the self-revelation model.[21]

E. Estimated Effect of School Average SAT Score by Parental Education

Finally, we explored whether the effect of college selectivity varied by average years of parental education.[22] The interaction term for school average SAT and years of parental education was negative for both cohorts, implying a higher payoff to attending a more selective school for students from more disadvantaged family backgrounds (Table 8). For example, in the self-revelation model for the 1989 cohort, our results suggest that attending a college with a 200-point higher average SAT score would lead to 5.2 percent higher earnings in 2007 for those

---

[20] See Fryer and Greenstone (2010) for estimates of the effect of HBCUs on earnings.

[21] Because our sample only includes two HBCUs, it is difficult to examine more explicitly why our results were sensitive to the inclusion of HBCUs in 1976 but not 1989.

[22] Parental education was equal to the average of the mother's and father's education; if data were missing for one parent, the average was set equal to the years of education for the parent with available data. The 13 students in the 1989 cohort and 22 students in the 1976 cohort that were missing education data for both parents were excluded from these regressions.

with average parental education of 12 years (equivalent to graduating from high school); however, for those whose parents averaged 16 years of education (approximately equivalent to college graduates), there was virtually no return to attending a more selective college. Similar to Dale and Krueger (2002), we also found a negative interaction between predicted parental income and school average SAT score, though the interaction term was generally not statistically significant.

VI. Conclusion

Consistent with the past literature, we find a positive and significant effect of college selectivity during a student's prime working years in regression models that do not adjust for unobserved student quality for cohorts that entered college in 1976 and 1989 using administrative earnings data from the SSA's Detailed Earnings Records. Based on these same regression specifications, we also find that the effect of college selectivity increases over the course of a student's career. However, after we partially adjust for unobserved student characteristics (by controlling for the average SAT score of the colleges students applied to) in our "self-revelation" model, the effect of college selectivity falls dramatically. For the 1976 cohort, the effect of school SAT score for the full sample is indistinguishable from zero in the self-revelation model. Similarly, the effects of other college characteristics (the Barron's index and net tuition) are substantial in regressions that control for commonly observed student characteristics but small and not statistically distinguishable from zero in the self-revelation model.

There were noteworthy exceptions for subgroups. First, for the 1989 cohort, the estimates indicate the effect of attending a school with a higher average SAT score is positive for black and Hispanic students, even in the selection-adjusted model. Second, our results suggest that students from disadvantaged family backgrounds (in terms of educational attainment) experience a greater

benefit from attending a college with a higher average SAT score than do those from more advantaged family backgrounds. For example, for the 1989 cohort, our estimates from the selection-adjusted model imply that the effect of attending a college with a higher average SAT score is positive for students whose parents had an average of fewer than 16 years of schooling; however, the effect of attending a more selective college was zero (or even negative) for students whose parents averaged 16 or more years of education. One possible explanation for this pattern is that although most students who apply to selective colleges may be able to rely on their families and friends to provide job-networking opportunities, networking opportunities that become available from attending a selective college may be particularly valuable for black and Hispanic students and for students from less educated families.

Contrary to expectations, our estimates do not suggest that the effects of college characteristics (within the set of C&B schools) increased for students who entered college more recently; estimates for the 1976 and 1989 cohort are similar when we compare the effects for each cohort at a similar stage relative to college entry (approximately 18 to 19 years after the students entered college). Specifically, for both cohorts, attending a college with a 100-point higher SAT score led to students receiving about 6 percent higher earnings (in 1995 and 2007, respectively) according our basic model; for both cohorts, this effect was close to zero in our selection-adjusted model.

Our findings have several caveats. First, the analysis does not pertain to a nationally representative sample of schools because the sample is derived from 27 colleges and universities in the C&B data set, the majority of which are very selective. However, estimates of the effects of school selectivity based on the C&B data set were similar to—indeed, slightly higher than—those based on a nationally representative data set, the NLS-72 (see Dale and Krueger 2002). In addition,

Dale and Krueger (2002) found an insignificant payoff to attending more selective schools when they used the NLS to estimate the self-revelation model. Thus, although the results reported in this paper are based on students that mainly attended moderately selective or very selective schools, it is not clear that we would have obtained different results from a nationally representative data set.

Second, the estimates from the selection-adjusted models are imprecise, especially for the 1989 cohort. Thus, even though the point estimates for the effect of a college characteristic are close to zero, the upper bound of the 95 percent confidence intervals for these estimates are sometimes sizeable. Also, our estimates are based on a single proxy for school quality, and therefore may be understated relative to estimates are based on multiple proxies for school quality as explained by Black and Smith (2006). Nonetheless, our results do imply that estimates that do not adjust for unobserved student characteristics are biased upward.

Finally, it is possible that our estimates are affected by students sorting into the colleges they attended based on their unobserved earnings potential. About 35 percent of the students in each cohort in our sample did not attend the most selective school to which they were admitted.[23] Our analysis indicates that students (especially those from the 1976 cohort) who were more likely to attend the most selective school to which they were admitted tended to have observable characteristics that are associated with higher earnings potential. If unobserved characteristics bear a similar relationship to college choice, then our already small estimates of the payoff from

[23] Hoxby (2009) mistakenly reports that only 10 percent of students in the C&B sample used in Dale and Krueger (2002) did not attend the most selective college to which they were admitted. However, similar to the results reported here, 38 percent of the students in the C&B sample used in Dale and Krueger (2002) did not attend the most selective college to which they were admitted.

attending a selective college would be biased upward. It is also possible that the benefit in terms of future earnings from attending a selective college varies across students and that students sort into college based on their perceived costs and benefits. Very selective colleges may attract not only students with very high family incomes (who can afford tuition) but also those with low family incomes (who receive financial aid). Conversely, students who expect a lucrative career because they intend to earn an MBA after college (for example) may sort into less selective undergraduate colleges. If students sort on the basis of their idiosyncratic return from attending a selective college, then equation (1) cannot be given a causal interpretation. However, if this is the case, then the typical student does not unambiguously benefit from attending the most selective college to which he or she was admitted. Rather, students need to think carefully about the fit between their abilities and interests, the attributes of the school they attend, and their career aspirations.

References

American Council on Education. 1973, 1983, 1992. *American Universities and Colleges*, 11th, 13th, 14th eds. Hawthorn, NY: Walter de Gruyter.

Angrist, Joshua, and Alan Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, vol. 3A, eds. Orley Ashenfelter and David Card, 1277–1366. Amsterdam: North Holland.

Barron's. *Barron's Profiles of American Colleges* (Great Neck, NY: Barron's Educational Series, 1978.)

Behrman, J. R., M. R. Rosenzweig, and P. Taubman. 1996. "College Choice and Wages: Estimates Using Data on Female Twins." *Review of Economics and Statistics* 78(4):672–85.

Black, Dan A., and Jeffrey A. Smith. 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence From Matching." *Journal of Econometrics* 121:99–124.

Black, Dan A., Kermit Daniel, and Jeffrey Smith. 2005. "University Quality and Wages in the United States." *German Economic Review* 6(3):415–43.

Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24 (3): 701–728.

Bound, J., and G. Johnson. 1992. "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations." *American Economic Review* 82:371–392.

Bowen, William G., and Derek Bok. 1998. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions.* Princeton, NJ: Princeton University Press.

Brewer, D., and R. Ehrenberg. 1996. "Does It Pay to Attend an Elite Private College? Evidence From the Senior Class of 1980." *Research in Labor Economics* 15:239–272.

Brewer, D., Eric Eide, and Ronald Ehrenberg. 1999. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *The Journal of Human Resources* 34(1):104–123.

College Division of Barron's Education Series (Ed.). 1992. *Barron's Profiles of American Colleges*, 19th ed. Hauppauge, NY: Barron's Educational Series, Inc.

Dale, Stacy Berg, and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4):1491–1528.

Fox, M.A., B.A. Connolly, and T.D. Snyder. 2005.*Youth Indicators 2005: Trends in the Well-Being of American Youth* (NCES 2005–050). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Fryer, R., and M. Greenstone. 2010. “The Changing Consequences of Attending Historically Black Colleges and Universities.” *American Economic Journal: Applied Economics* 2(1):116–48.

Grogger, J., and E. Eide. 1995. “Changes in College Skills and the Rise in the College Wage Premium.” *Journal of Human Resources* 30:280–310.

Hershbein, Brad. “Worker Signals Among New College Graduates: The Role of Selectivity and GPA.” Working Paper. The University of Michigan. October, 2011.

Hoekstra, Mark. 2009. “The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach.” *The Review of Economics and Statistics* 91(4): 717-724.

Hoxby, Caroline. 2009. “The Changing Selectivity of American Colleges.” *Journal of Economic Perspectives* 23(4):95–118.

Katz, Lawrence F., and Kevin M. Murphy. 1992. “Changes in Relative Wages, 1963–87: Supply and Demand Factors.” *Quarterly Journal of Economics* 107(February):35–78.

Lindahl, Lena, and Hakan Regner. 2005. “College Choice and Subsequent Earnings: Results Using Swedish Sibling Data.” *Scandinavian Journal of Economics* 107(3):437–57.

Long, Mark. 2008. “College Quality and Early Adult Outcomes.” *Economics of Education Review* 27(5):588–602.

Long, Mark. 2009. “Changes in the Returns to Education and College Quality. *Economics of Education Review* 29 (3) 338-347.

Loury, L. D., and D. Garman. 1995. “College Selectivity and Earnings.” *Journal of Labor Economics* 13(2):289–308.

Monks, James. 2000. “The Returns to Individual and College Characteristics Evidence from the National Longitudinal Survey of Youth.” *Economics of Education Review* 19(3):279–89.

Roy, A.D. “Some Thoughts on the Distribution of Earnings.” Oxford Economic Papers (New Series) 3 (1951): 135-146.

Van Der Klaauw. 1997. “A Regression Discontinuity Evaluation of the Effect of Financial Aid Offers on Enrollment.” C.V. Starr Center Research Report 97-10, New York University.

Table 1

Descriptive Statistics

	1976 cohort				1989 cohort			
	Full sample		Black and Hispanic		Full sample		Black and Hispanic	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Dependent variables: Earnings measures, 2007 dollars								
Log 2007 earnings	11.53	1.07	11.29	0.81	11.41	1.21	11.21	0.73
2007 annual earnings	183,411	711,803	119,861	242,161	139,698	359,255	98,472	107,852
Log (median 1993 through 1997 earnings)	11.30	0.847	10.78	1.02				
Log (median 1993 through 1997 earnings)	11.48	1.022	11.00	1.19				
Median of 1993 through 1997 earnings	106,638	148,615	70,760	77,178				
Median of 2003 through 2007 earnings	164,009	489,647	102,714	211,876				
Regression control variables								
Average school SAT/100	11.58	1.22	10.95	2.00	12.0	1.4	11.54	1.45
Average SAT score of schools applied to/100	11.40	1.20	10.73	1.83	11.9	1.4	11.33	1.33
Student SAT/100	11.61	1.89	9.46	2.15	12.1	2.6	10.24	2.20
Student SAT is missing	0.05	0.24	0.07	0.26	0.00	0.09	0.02	0.16
Log net tuition	7.66	0.55	7.69	0.47	8.95	0.73	8.97	0.41
Barron's index	3.34	1.20	3.32	1.19	4.19	1.13	3.74	1.21
Female	0.43	0.56	0.53	0.51	0.45	0.72	0.51	0.50
Black	0.06	0.27	0.88	0.33	0.08	0.40	0.74	0.44
Hispanic	0.01	0.10	0.12	0.33	0.03	0.24	0.26	0.44
Asian	0.02	0.16	0.00	0.00	0.08	0.40	0.00	0.00
Other race	0.04	0.22	0.00	0.00	0.00	0.10	0.00	0.00
High school GPA, 4-point scale	3.57	0.36	3.37	0.44	3.62	0.36	3.49	0.36
High school GPA missing	0.37	0.54	0.40	0.50	0.61	0.71	0.59	0.49
Predicted parental income	9.98	0.39	9.70	0.42	11.05	0.54	10.82	0.45
Student athlete	0.07	0.29	0.06	0.23	0.08	0.39	0.06	0.23
Applications submitted								
0 additional	0.37	0.48	0.38	0.49	0.29	0.45	0.31	0.46
1 additional	0.22	0.42	0.24	0.44	0.20	0.58	0.21	0.41

2 additional	0.21	0.41	0.21	0.41	0.22	0.60	0.22	0.42
3 additional	0.15	0.36	0.14	0.35	0.23	0.61	0.21	0.41
4 additional	0.04	0.22	0.03	0.17	0.07	0.38	0.05	0.21
Sample size (unweighted)	12,075		1,167		6,479		1,508	

Source: Data from the C&B Survey and Detailed Earnings Records from the Social Security Administration.

Notes: Means are weighted to make the sample representative of all students in the C&B institutions. Means for the regression control variables for the 1976 cohort reflect the sample used in the 1993–1997 earnings regression (reported in Table 3). Means for the 1989 cohort reflect the sample used in the 2007 earnings regression (as reported in Table 7). Earnings means exclude those with annual earnings below the minimum wage threshold (equivalent to $13,822 in 2007 dollars).

Std = standard deviation.

Table 2

College Application Patterns Among Students Attending College and Beyond Schools

	1976 cohort		1989 cohort	
	Full sample (%)	Blacks and Hispanics (%)	Full sample (%)	Blacks and Hispanics (%)
Percentage of students submitting at least one additional application	64.3	64.1	71.3	70.1
Among those submitting at least one additional application				
Applied to a school with a higher average SAT score than the school attended	53.7	49.0	55.0	48.4
Percentage accepted to at least one additional school	87.9	94.0	88.3	92.5
Among those accepted to at least one additional school				
Percentage accepted to a school with a higher average SAT score than the one they attended	35.0	40.2	36.1	40.3
Percentage accepted to a school with an SAT score at least 40 points higher than the average SAT score than the one they attended	22.6	28.2	23.0	27.4
Sample size (unweighted)	17,223	1,411	8,830	2,016

Source: C&B Survey.

Notes: Sample includes all survey respondents from C&B schools participating in this study. Means are weighted to make the sample representative of all students attending the C&B schools participating in this study.

Table 3

Comparing Parameter Estimates of the Effect of College Average SAT Score on Earnings Using C&B and SSA Data, 1976 Cohort

	C&B sample[a]		Merged C&B and SSA sample[b]									
	Log 1995 C&B earnings		Log 1995 C&B earnings		Log 1995 SSA earnings (topcoded)		Log 1995 SSA earnings (not topcoded)		Log (median of 1993 to 1997 earnings), SSA data		Log (median of 1993 to 1997 earnings), SSA data	
	1	2	3	4	5	6	7	8	9	10	11	12
	Basic	Self – revelation	Basic	Self - revelation	Basic	Self – revelation	Basic	Self - revelation	Basic	Self – revelation	Basic	Self – revelation
Parameter estimate for school SAT/100	0.076	-0.001	0.068	-0.007	0.048	-0.021	0.058	-0.015	0.059	-0.025	0.061	-0.023
	(.008)	(.012)	(.007)	(.012)	(.009)	(.014)	(.009)	(.015)	(.008)	(.012)	(.007)	(.012)
	{.016}	{.018}	{.014}	{.018}	{.016}	{.018}	{.017}	{.016}	{.012}	{.013}	{.013}	{.014}
N	14,238		10,886		10,886		10,886		11,932		12,075	
Sample restriction	Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Full-time workers (according to C&B survey)		Median earnings greater than zero (SSA data)		Median earnings greater than $13,822 in 2007 dollars (SSA data)	

Source: C&B Survey and SSA's Detailed Earnings Records.

Notes: Each cell corresponds to an estimate drawn from a different weighted least squares regression that controlled for race, student SAT score, high school GPA, dummies for whether high school GPA or SAT score is missing, predicted parental income, and student athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and for the number of applications the student submitted. Two sets of standard errors are reported, one in parentheses and one in brackets; those in brackets are robust to correlated errors among students who attended the same colleges. The top earnings category for the C&B data (more than $200,000) was topcoded at $242,662; SSA earnings data in columns 5 and 6 were topcoded in the same way.

[a]Sample includes survey respondents from the 30 C&B institutions analyzed in Dale and Krueger (2002).

[b]Sample includes survey respondents from the 27 C&B institutions participating in this study that were matched to SSA data.

Table 4

Effect of School SAT Score/100 on Having Earnings Greater Than Minimum Threshold

	1976 men		1976 women		1976 pooled men and women		1989 pooled men and women	
	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation
Had earnings greater than minimum wage threshold ($13,822 in 2007 dollars)								
Parameter estimate for school SAT score/100	-0.062	-0.047	-0.067	-0.034	-0.061	-0.049	-0.093	-0.085
	{.039}	{.038}	{.027}	{.035}	{.026}	{.034}	{.072}	{.074}
Had non-zero earnings								
Parameter estimate for school SAT score/100	-0.059	-0.067	-0.069	-0.031	-0.062	-0.048	-0.097	-0.088
	{ .031}	{.041}	{.032}	{.040}	{.028}	{.038}	{.077}	{.080}
N	8,304		8,897		17,202			8,830

Source: C&B Survey and SSA's Detailed Earnings Records.

Notes: Each cell corresponds to an estimate drawn from a different weighted least squares regression where the dependent variable was whether the student had earnings that were greater than the minimum wage threshold (in the top panel) or greater than zero (in the bottom panel). The regressions each controlled for race, student SAT score, high school grade point average, dummies for whether high school grade point average or SAT score is missing, predicted parental income, and student athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and for the number of applications the student submitted. Standard errors are reported in brackets and are robust to correlated errors among students who attended the same colleges.

Table 5

Effect of School SAT Score/100 on Earnings, 1976 Cohort

	Men		Women		Men and women pooled	
	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation
Effect on log (median of 1983 through 1987 earnings)						
Parameter estimate for school SAT/100	0.011	-0.004	-0.007	-0.037	0.004	-0.017
	(.007)	(.011)	(.007)	(.012)	(.005)	(.008)
	{.012}	{.019}	{.012}	{.023}	{.011}	{.014}
N	6,294	6,294	5,690	5,690	11,984	11,984
Effect on log (median of 1988 earnings through 1992 earnings)						
Parameter estimate for school SAT/100	0.054	-0.001	0.031	-0.034	0.045	-0.014
	(.009)	(.013)	(.009)	(.015)	(.006)	(.010)
	{.014}	{.016}	{.014}	{.019}	{.012}	{.013}
N	6,911	6,911	6,294	6,294	12,407	12,407
Effect on log (median of 1993 earnings through 1997 earnings)						
Parameter estimate for school SAT/100	0.080	0.001	0.034	-0.059	0.061	-0.023
	(.010)	(.016)	(.010)	(.018)	(.007)	(.012)
	{.013}	{.015}	{.012}	{.016}	{.013}	{.014}
N	6,896	6,896	5,179	5,179	12,075	12,075
Effect on log (median of 1998 earnings through 2002 earnings)						
Parameter estimate for school SAT/100	0.087	0.002	0.042	-0.069	0.070	-0.024
	(.012)	(.018)	(.012)	(.020)	(.008)	(.013)
	{.015}	{.022}	{.010}	{.016}	{.012}	{.019}
N	6,869	6,869	5,195	5,195	12,064	12,064
Effect on log (median of 2003 earnings through 2007 earnings)						
Parameter estimate for school SAT/100	0.083	0.006	0.057	-0.035	0.074	-0.008
	(.013)	(.020)	(.012)	(.021)	(.009)	(.014)
	{.015}	{.027}	{.013}	{.018}	{.014}	{.018}
N	6,650		5,244		11,894	

Source: C&B Survey and Detailed Earnings Records from the Social Security Administration.

Notes: Each cell represents parameter estimates from a separate weighted least squares regression. Each model controls for race, gender (in the pooled model only), predicted parental income, student's SAT score, student's high school grade point average, whether the student was a college athlete, and dummies indicating when high school GPA or SAT score were missing; the self-revelation model also controls for the average SAT score of the schools to which the student applied and for the number of applications the student submitted. Two sets of standard errors are reported, one in parentheses and one in brackets. Standard errors in brackets are

robust to correlated errors among students who attended the same institution. Individuals are excluded if the median of annual earnings over the five-year interval was less than $13,822 in 2007 dollars.

Table 6

Effect of College Characteristics on Earnings, 1976 Cohort of Men and Women

	College characteristic: Log net tuition		College characteristic: Barron's index	
	Basic	Self-revelation	Basic	Self-revelation
Effect on log (median of 1983 through 1987 annual earnings)				
Parameter estimate for	0.014	-0.007	0.010	0.001
school quality measure	(.010)	(.013)	(.005)	(.013)
N = 11,984	{.024}	{.027}	{.012}	{.015}
Effect on log (median of 1988 through 1992 annual earnings)				
Parameter estimate for	0.092	0.012	0.055	0.020
school quality measure	(.012)	(.016)	(.006)	(.017)
N = 12,407	{.028}	{.028}	{.011}	{.015}
Effect on log (median of 1993 through 1997 annual earnings)				
Parameter estimate for	0.124	0.013	0.071	0.017
school quality measure	(.015)	(.019)	(.007)	(.010)
N = 12,075	{.030}	{.038}	{.009}	{.015}
Effect on log (median of 1998 through 2002 annual earnings)				
Parameter estimate for	0.140	0.017	0.077	0.014
school quality measure	(.012)	(.017)	(.008)	(.012)
N = 12,064	{.026}	{.034}	{.008}	{.019}
Effect on log (median of 2003 through 2007 annual earnings)				
Parameter estimate for	0.143	0.026	0.080	0.023
school quality measure	(.018)	(.023)	(.009)	(.012)
N = 11,894	{.032}	{.039}	{.010}	{.017}

Source: C&B Survey and Detailed Earnings Records from the Social Security Administration.

Notes: Each cell represents parameter estimates from a separate weighted least squares regression. Both the basic and self-revelation models control for race, sex, predicted parental income, student's SAT score, a dummy indicating if student SAT score was missing, student's high school grade point average, a dummy indicating if high school grade point average was missing, and whether the student was a college athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and dummies for the number of applications the student submitted. Weights were used to make the sample representative of students at C&B schools. Two sets of standard errors are reported, one in parentheses and in brackets. Standard errors in brackets are robust to correlated errors among students who attended the same institution. The Barron's measure is coded as a continuous measure, ranging from 2 (Competitive colleges) to 5 (Most Competitive colleges) for our sample. Individuals are excluded if their median annual earnings over the five-year interval were less than $13,822 in 2007 dollars.

Table 7

Effect of College Characteristics on 2007 Earnings, 1989 Cohort of Men and Women

	College characteristic					
	School SAT score/100		Log net tuition		Barron's index	
	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation
Parameter estimate for	0.056	-0.008	-0.011	-0.108	0.069	-0.002
effect of quality measure	(.014)	(.019)	(.025)	(.028)	(.017)	(.022)
on log 2007 earnings	{.031}	{.034}	{.062}	{.070}	{.038}	{.042}
Sample size	6,479		6,479		6,479	

Source: C&B data and Detailed Earnings Records from the Social Security Administration.

Notes: Parameter estimates drawn from a weighted least squares regression. Each cell represents a different regression. Both the basic and self-revelation models control for race, sex, predicted parental income, student's SAT score, a dummy indicating if student SAT score was missing, student's high school grade point average, a dummy indicating if high school grade point average was missing, and whether the student was a college athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and dummies for the number of applications the student submitted. Weights were used to make the sample representative of students at C&B schools. Two sets of standard errors are reported, one in parentheses and in brackets. Standard errors in brackets are robust to correlated errors among students who attended the same institution. The Barron's measure is coded as a continuous measure, ranging from 2 (Competitive colleges) to 5 (Most Competitive colleges) for our sample. Individuals were excluded if their annual earnings were less than $13,822 in 2007 dollars.

Table 8

Effect of School Characteristics on 2007 Earnings (Black and Hispanic Students Only, 1989 Cohort)

	School SAT score/100		Log net tuition		Barron's index	
Dependent variable	Basic	Self-revelation	Basic	Self-revelation	Basic	Self-revelation
All black and Hispanic students						
Parameter estimate for	0.067	0.076	0.173	0.138	0.063	0.049
effect of quality measure	(.019)	(.032)	(.056)	(.071)	(.022)	(.036)
on log 2007 earnings	{.028}	{.042}	{.076}	{ .092}	{.033}	{.046}
Sample size	1,508		1,508		1,508	
All black and Hispanic students, excluding historically black colleges and universities						
Parameter estimate for	0.122	0.120	0.187	0.116	0.158	0.143
effect of quality measure	(.030)	(.042)	(.064)	(.079)	(.040)	(.053)
on log 2007 earnings	{.035}	{.056}	{.081}	{.101}	{.038}	{.051}
Sample size	995		995		995	

Source: C&B Survey and Social Security Administration's Detailed Earnings Records.

Notes: Parameter estimates drawn from weighted least squares regression models. Weights were used to make the sample representative of the population of students at C&B schools. Both the basic and self-revelation models control for race, sex, predicted parental income, student SAT score, student high school grade point average, dummy variables indicating if high school grade point average or student SAT score was missing, and whether the student was a college athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and dummies for the number of applications the student submitted. Two sets of standard errors are reported, one in parentheses and one in brackets. Standard errors in brackets are robust to correlated errors among students who attended the same institution. The Barron's measure is coded as a continuous measure, ranging from 2 (Competitive colleges) to 5 (Most Competitive colleges) for our sample. Individuals are excluded if they earned less than $13,822 in 2007.

Table 9

Parameter Estimates from Earnings Regressions, Allowing the Effect of Average School SAT to Vary by Parental Education

	Parameter estimates			
	1976 Cohort		1989 Cohort	
Variable	Basic	Self-revelation	Basic	Self-revelation
School SAT score/100	0.126	0.041	0.206	0.117
	(.035)	(.036)	(.073)	(.075)
	{.058}	{.055}	{.110}	{.112}
Average years of parental education	0.066	0.063	0.128	0.106
	(.027)	(.027)	(.053)	(.053)
	{.041}	{.038}	{.097}	{.097}
Years of parental education * school SAT score/100	-0.004	-0.004	-0.009	-0.008
	(.002)	(.002)	(.004)	(.004)
	{.0037}	{.003}	{.008}	{.008}
Effect of a 200-point increase in school SAT score if				
Average years of parental education = 12 (equivalent to high school graduate)	0.148	-0.025	0.193	0.052
Average years of parental education = 16 (approximately equivalent to college graduate)	0.113	-0.060	0.120	-0.009
Average years of parental education = 19 (approximately equivalent to graduate degree)	0.087	-0.087	0.065	-0.055
Sample size (unweighted)	12,053	12,053	6,466	6,466

Source: C&B data and Detailed Earnings Records from the Social Security Administration.

Notes: Parameter estimates drawn from a weighted least squares regression. In addition to the variables listed in the first column of the table, both the basic and self-revelation models control for race, gender, student's SAT score, a dummy indicating missing student SAT score, student's high school GPA, a dummy indicating if high school GPA was missing, and a dummy for college athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and for the number of applications the student submitted. Two sets of standard errors are reported, one in parentheses and one in brackets. Standard errors in brackets are robust to correlated errors among students who attended the same institution. The dependent variable for the 1989 cohort is the log 2007 earnings and for the 1976 cohort is the log (median of 1993 annual earnings through log 1997 annual earnings). Individuals are excluded if their annual earnings (for the 1989 cohort) or median annual earnings (for the 1976 cohort) were less than $13,822 in 2007 dollars. The parental education measure is the average of the mother's and father's education; if data was missing for one parent, the average was set equal to the years

of education of the parent with data. Individuals with no parental education data for either parent are excluded. The average (standard deviation) of years of parental education was 15.23 (2.70) for the 1976 cohort and 16.18 (3.24) for the 1989 cohort.

Appendix

Table A1

Distribution of School Characteristics for College and Beyond Schools

College characteristic	Number of schools	Number of students (unweighted)	Number of students (weighted)
1976 cohort			
Average SAT score			
1250 and greater	6	2,962	2,962
1175–1250	13	7,999	7,999
< 1175	8	6,262	10,219
1976 net tuition			
> $3,500	10	5,497	5,497
$2,000–$3,500	12	7,164	7,164
< $2,000	5	4,562	8,519
1978 Barron's index			
Most competitive	11	5,503	5,503
Highly competitive	7	4,833	4,833
Very competitive or competitive	9	6,887	10,844
1989 cohort			
Average SAT score			
1250 and greater	7	3,604	5,944
1175–1250	6	2,857	4,220
< 1175	5	2,369	8,488
1990 net tuition			
> $12,000	6	2,971	4,116
$10,000–$12,000	7	3,490	6,048
< $10,000	5	2,369	8,488
1992 Barron's index			
Most competitive	10	4,931	7,743
Highly competitive	5	2,651	7,909
Very competitive or competitive	3	1,248	3,001

Source: C&B Survey; American Council on Education (1973, 1983, 1992) (for net tuition); College Division of Barron's Education Series (1992).

Notes: The C&B Survey in 1989 included only a subset of schools included in the 1976 survey. For both cohorts, the sample is weighted to make the sample representative of the population of students attending the C&B institutions. Net tuition was calculated by subtracting the average aid awarded to undergraduates from the sticker price tuition.

Table A2
Characteristics of Colleges in Sample

	1976 cohort			1989 cohort		
Institution	1976 average SAT score	1976 net tuition ($)	1978 Barron's index	1989 average SAT score	1990 net tuition ($)	1992 Barron's index
Barnard College	1210	3,530	4			
Bryn Mawr College	1370	3,171	5	1285	11,905	5
Columbia University	1330	3,591	5			
Duke University	1226	3,052	4	1285	12,219	5
Emory University	1150	3,237	3			
Georgetown University	1225	3,304	4	1235	12,141	5
Hamilton College	1246	3,529	4			
Miami University (Ohio)	1073	1,304	2	1154	3,018	4
Morehouse	830	1,549	2	998	5,601	3
Northwestern University	1240	3,676	4			
Oberlin College	1227	3,441	4	1216	15,738	5
Pennsylvania State University	1038	1,062	2	1083	5,426	3
Princeton University	1308	3,613	5	1255	11,800	5
Smith College	1210	3,539	5			
Stanford University	1270	3,658	5	1351	11,127	5
Swarthmore College	1340	3,122	5			
Tufts University	1200	3,853	4			
Tulane University	1080	3,269	3			
University of Michigan	1110	1,517	3	1164	6,595	4
University of Notre Dame	1200	3,216	4			
Vanderbilt University	1162	3,155	3	1177	11,030	4
Washington University	1180	3,245	3	1201	11,267	5
Wellesley College	1220	3,312	5	1240	11,358	5
Wesleyan University	1260	3,368	5	1284	12,688	5
Williams College	1255	3,541	5	1329	13,249	5
Xavier	710	1,326	4	966	5,391	2
Yale University	1360	3,744	5	1379	13,366	5

Source: C&B Survey, American Council on Education (1973, 1983, 1992) (for net tuition); College Division of Barron's Education Series (1992).

Notes: The C&B Survey in 1989 included only a subset of schools included in the 1976 survey; the college characteristic is not reported for those colleges that were not included in the 1989 survey. For both cohorts, the sample is weighted to make the sample representative of the population of students attending the C&B institutions. Net tuition was calculated by subtracting the average aid awarded to undergraduates from the sticker price tuition. The Barron's measure is coded as a continuous measure, ranging from 2 (Competitive colleges) to 5 (Most Competitive colleges) for our sample.

Table A3

Full Set of Parameter Estimates for Selected Log of Earnings Regressions

	1976 Cohort		1989 Cohort	
Variable	Basic	Self-revelation	Basic	Self-revelation
School SAT score/100	0.061	-0.023	0.056	-0.008
	(.013)	(.014)	(.031)	(.034)
Student SAT score/100	0.022	0.014	0.047	0.033
	(.005)	(.005)	(.008)	(.008)
Student SAT missing	-0.141	-0.122	-0.262	-0.217
	(.030)	(.030)	(.160)	(.160)
Female	-0.479	-0.469	-0.410	-0.412
	(.013)	(.013)	(.020)	(.020)
Black	-0.028	-0.037	0.036	0.022
	(.029)	(.029)	(.040)	(.040)
Hispanic	-0.063	-0.077	-0.060	-0.074
	(.069)	(.069)	(.059)	(.040)
Asian	0.171	0.151	0.154	0.139
	(.046)	(.046)	(.036)	(.036)
Other race	-0.088	-0.101	-0.363	-0.344
	(.034)	(.034)	(.143)	(.143)
High school GPA	0.218	0.216	0.194	0.188
	(.021)	(.021)	(.042)	(.042)
High school GPA missing	0.015	0.013	0.094	0.092
	(.014)	(.014)	(.021)	(.021)
Predicted parental income	0.161	0.140	0.137	0.117
	(.019)	(.017)	(.029)	(.029)
Athlete	0.124	0.123	0.135	0.092
	(.025)	(.037)	(.037)	(.020)
Average SAT score/100 of schools applied to		0.100		0.099
		(.012)		(.014)
One additional application		0.062		0.029
		(.017)		(.029)
Two additional applications		0.057		0.053
		(.018)		(.028)
Three additional applications		0.073		0.084
		(.020)		(.028)
Four additional applications		0.085		0.098
		(.034)		(.041)
R-squared	0.147	0.153	0.122	0.126
Sample size (unweighted)	12,075		6,479	

Source: C&B Survey and Detailed Earnings Records from the Social Security Administration.

Notes: Parameter estimates drawn from weighted least squares regression models where the dependent variable is log 2007 earnings for the 1989 cohort and log (median of 1983 through 1987 annual earnings) for the 1976 cohort. Standard errors are in parentheses and are robust to correlated errors among students who attended the

same institution. Individuals are excluded if annual earnings (for the 1989 cohort) or the median of annual earnings (for the 1976 cohort) were less than $13,822 in 2007 dollars. Weights were used to make the sample representative of the population of students at C&B schools.

Table A4

Relationship Between Student Characteristics and Average SAT Score/100 of College Attended

	Correlation with college average SAT score		Parameter estimate for effect of student characteristic on college average SAT score of school attended	
	1976 cohort	1989 cohort	1976 cohort	1989 cohort
Predicted parental income	0.182 < 0.001	0.278 < 0.001	-0.020 (.013)	0.049 (.016)
Student SAT score/100	0.511 < 0.001	0.579 < 0.001	0.060 (.060)	0.049 (.004)
High school grade point average	0.265 < 0.001	0.200 < 0.001	0.114 (0.014)	-0.052 (.024)
Female	0.016 0.0345	0.023 0.033	0.084 (.009)	0.067 (.001)
Black	-.184 < 0.001	-0.232 < 0.001	-0.066 (.019)	-0.086 (.022)
Hispanic	0.062 < .001	0.084 < 0.001	0.305 (.043)	0.192 (.033)
Asian	0.090 < 0.001	0.160 < 0.001	0.113 (.030)	0.036 (.020)
Other race	-0.041 < .001	0.029 0.006	-0.132 (.022)	0.260 (.076)

Source: C&B Survey.

Notes: The first two columns show the correlations between student characteristics and the average SAT score of the college they attended. The third and fourth columns show parameter estimates for the student characteristic shown in the left margin; each parameter estimate is drawn from a separate weighted least squares regression model that estimates the effect of the student characteristic on the average SAT score of the college attended, after controlling for the average SAT score of the schools the student applied to. Standard errors are in parentheses. Weights were used to make the sample representative of the population of students at C&B schools.

Table A5

Effect of Barron's Categories on Log of Earnings

	Effect on 1995 earnings, 1976 cohort		Effect on 2007 earnings,1976 cohort		Effect on 2007 earnings, 1989 cohort
	Most competitive college	Highly competitive college	Most competitive college	Highly competitive college	Most competitive college
Basic					
	0.154	0.101	0.190	0.107	.101
	(.020)	(.017)	(.025)	(.-021)	(.025)
	{.031}	{.051}	{ .033}	{ .051}	{ .066}
Self-revelation	0.013	0.004	.027	-0.022	.018
	(.026)	(.021)	(.032)	(.025)	(.030)
	{.044}	{ .056}	{ .038}	{ .056}	{ .087}
Sample size	11,932		11,725		6,479

Source: C&B Survey and Detailed Earnings Records from the Social Security Administration.

Notes: Each cell represents parameter estimates from a separate weighted least squares regression. Both the basic and self-revelation models control for race, sex, predicted parental income, student's SAT score, a dummy indicating if student SAT score was missing, student's high school grade point average, a dummy indicating if high school grade point average was missing, and whether the student was a college athlete; the self-revelation model also controls for the average SAT score of the schools to which the student applied and dummies for the number of applications the student submitted. The omitted categories include the Competitive and Very Competitive categories for the 1976 cohort and the Competitive, Very Competitive, and Highly Competitive categories for the 1989 cohort. Weights were used to make the sample representative of students at C&B schools. Two sets of standard errors are reported, one in parentheses and in brackets. Standard errors in brackets are robust to correlated errors among students who attended the same institution. Individuals with annual earnings less than $13,822 in 2007 dollars were excluded.