Ova and Out: Using Twins to Estimate the Educational Returns to Attending a Selective College

Jonathan Smith

The College Board – Advocacy & Policy Center

June 2013

Abstract: Research has shown that attending a relatively selective four-year college over a less selective alternative is positively related to bachelor's degree completion. This paper revisits that question with a novel dataset of over 11,000 sets of twins in the United States and information on colleges to which they apply, enroll, and potentially graduate. I show that a student's probability of bachelor's degree completion within four years increases by 5 percentage points by choosing an institution with a median SAT score 100 points higher than the alternative. Moreover, the estimated magnitude of impact is insensitive to several methodologies, including OLS, twin fixed effects, and controlling for the application portfolio. This suggests that in certain contexts, sources of bias perceived as barriers to obtaining causal estimates of the returns to college selectivity, such as unobserved family characteristics and student aspiration, may be of little concern.

JEL: I2, I23

Keywords: Higher Education; Returns to Selectivity; College Completion

Contact information: 1233 20th Street, NW, Suite 600, Washington, DC 20036. Email: jsmith@collegeboard.org. This research does not reflect the views of the College Board.

1. Introduction

Research shows that there is a strong positive relationship between college selectivity and graduation rates (Bowen and Bok 1998; Kane 1998; Alon and Tienda 2005; Horn and Carroll 2006; Long 2008; and Bowen et al. 2009). Similarly, there is a strong relationship between college selectivity and longer term outcomes such as graduate school attendance (Eide, Brewer, and Ehrenberg, 1998; Brand and Halaby, 2006; Zhang; 2005) and wages (James et al. 1989; Loury and Garman 1995; Behrman, Rosensweig, and Taubman 1996; Daniel, Black, and Smith 1997; Hoxby 1998; Kane 1998; Brewer, Eide, and Ehrenberg 1999; and Monks 2000; Long 2008). Identifying causal relationships has been more challenging because selection on student unobservables is likely to bias estimates of the return to selectivity. However, researchers, typically looking at the effect on future wages, have developed compelling identification strategies to overcome the bias. For example, Dale and Krueger (2002, 2011) match students who have similar college application portfolios and acceptances to a highly selective set of institutions, arguing that these students are similar on unobservables. They find that college quality has no impact on future wages.[1] This analysis is replicated and confirmed with a broader set of institutions in Long (2008),

[1] There is a positive impact for minorities and when using net tuition as a measure of quality.

who finds that there is a positive impact on the probability of graduating college.[2] Hoekstra (2009) and Saavedra (2008) use a regression discontinuity approach based on admission cutoffs at a single flagship university and in Columbia, respectively, and find that matriculates at the more selective universities earn approximately 20 percent more than applicants with scores below the admission threshold. Finally, Cohodes and Goodman (2012) use a regression discontinuity that exploits cut scores for merit scholarships at public universities in Massachusetts. They find that high achieving students who are compelled to enroll in a less selective (public) university than their peers have a 40 percent lower probability of graduating on time.

The aforementioned research has made great progress in controlling for selection bias. However, without the availability of a discontinuity to exploit, additional unobservable family characteristics are still likely to exist in most of the analyses and potentially bias estimates, as noted in Lindahl and Regner (2005). For example, parents may insist that their children attend a selective institution and graduate, regardless of parental income and education or other typical family-level observables. To address this, several researchers have addressed the selection bias related to family background by using twins (Ashenfelter and Krueger 1994; Rouse 1999).[3] In a related paper, Behrman,

---

[2] Long (2008) not only uses the Dale and Krueger (2002) method, but also uses propensity score methods from Black and Smith (2004) and an instrumental variables method.

[3] There is a substantial literature on issues that potentially arise when using within twin variation. These are discussed in more detail in Section 3.2.4.

Rosensweig, and Taubman (1996) use female twin data and find that wages are affected by college quality and years of schooling but stop short of suggesting that college quality directly impacts years of schooling.

This paper estimates the effect of attending a relatively selective college on the probability of graduating by using a novel set of twins and a rich set of information on students' application portfolios. Using College Board data, I identify 11,008 sets of twins, who both take the SATs and enroll in a four-year college, by matching students in the same high school, with the same last name, address, and date of birth. This sample size dwarfs previous twin studies, and therefore, I am able to examine heterogeneity in outcomes and to obtain precise estimates. Getting these precise and nuanced estimates is important because from the students' perspective, college is a large time, financial, and human capital investment and they deserve to know the investment's expected return and whether choosing one college over another has higher expected returns. From the policymakers' perspective, college is heavily subsidized and understanding the effect of college selectivity can help allocate resources more efficiently.

As a starting point, I regress whether a student graduates in four years on college selectivity, as measured by the median SAT score of enrollees, controlling for student characteristics and achievement measures and parent characteristics.[4] I find that

---

[4] Controls include student SAT, high schools GPA, whether participates in AP, number of SAT2's taken, ethnicity, first language, parental income and education, state residence, local unemployment and education attainment, and cohort.

attending a college that has a median SAT 100 points above the alternative is associated with a 5.8 percentage point increase in graduation probability.

Following Dale and Krueger (2002), I include controls for the number of applications and quality of colleges in the portfolio because there may exist unobservable differences among students who apply to different sets of colleges, such as aspiration or ability, which may bias estimates. Including these controls reduces the graduation impact estimate slightly to 5.2 percentage points per 100 point median SAT increase. I also match students on application portfolios so as to include portfolio fixed effects, but results are largely unchanged.[5] The robustness of these results, even when including application portfolio controls, is inconsistent with Long (2008), whose OLS estimate of the effect of selectivity on graduation rates is cut in half and not statistically different than zero.

Next, I estimate a twin fixed effects model and find, again, an estimate of 5.2 percentage point improvement in the probability of graduating when enrolling in a school with a median SAT score 100 points greater than the alternative. This result is consistent with previous research on twins, which suggests that controlling for the selection on unobservable family characteristics reduces the magnitude of estimates, but usually by very little (Ashenfelter and Krueger 1994; Rouse 1999). Finally, I combine the twin fixed effects model and include application portfolio controls and find that the estimate is

---

[5] Dale and Krueger (2002) match on application portfolio and acceptances. The costs and benefits of my approach are discussed in the empirical strategy.

largely unchanged- 4.8 percentage point increase in four-year graduation probability per 100 SAT point increase in median SAT.

The large sample size affords me the opportunity to test for nonlinear effects, which do exist: the largest gains in graduation probability occur when choosing moderately selective colleges over less selective colleges, rather than highly selective colleges over moderately selective colleges. Enrolling in a moderately selective college (median SAT between 1100 and 1199) has almost a 10 percentage point graduation advantage relative to enrolling in a less selective college (median SAT below 1100) whereas enrolling in a highly selective college (median SAT above 1199) has an additional 5 percentage point graduation advantage relative to enrolling in a moderately selective college (and a 15 percentage point advantage over less selective colleges). There are few differences in graduation rates within finer gradations of less selective colleges and highly selective colleges.

Next, I test for heterogeneous effects. I find that the relationship between institutional selectivity and four year graduation probability is nearly twice as large for males compared to females. Compared to White students, Black, Hispanic and Asian students are less impacted, in terms of graduation rates, by institutional selectivity. The relationship between institutional selectivity and graduation probability is also larger for students from suburban high schools, compared to their urban and rural counterparts.

Finally, I find evidence that undermatching does reduce a student's probability of graduating whereas overmatching has no pronounced effect.[6]

2. Data

2.1. General Data

The College Board data consists of all high school students who take the SAT. The data include the student's SAT score and performance on other College Board products, including the SAT2s and Advanced Placement (AP) tests. The data also include where students send their SAT scores (Score Sends), which is often required when applying to college.[7] Over the sample period, students have four free Score Sends to colleges, which are only free for a few days after the test taking day. After the test, each Score Send to a college costs $10.[8]

Though I do not know to which colleges students apply, I use Score Sends as a proxy for application, which has been shown to be a good measure of applications (Card and Krueger, 2005) and not unique to this paper (Pallais 2012). As an alternative

---

[6] Undermatching is when high ability students enroll in relatively unselective schools. Overmatching is when low ability students enroll in relatively selective schools.

[7] Approximately 65 percent of four year colleges require or recommend SAT or ACT scores and even more colleges use it for placement purposes. This sample of colleges only includes those that report median SAT of enrollees, which are typically more selective colleges who require the SAT from applicants.

[8] Low-income students can qualify to take the SAT for free and get free Score Sends.

interpretation, colleges that receive scores from students are in a student's choice set, which is an equally important metric.[9]

Finally, students fill out a questionnaire when they register for the SAT that provides demographic information (family income, parental education, race/ethnicity, gender, citizenship, and first language) as well as basic background information (name, address, date of birth, high school enrolled, and GPA). The student's high school was linked to the Common Core of Data, which provided school urbanicity.

In the summer of 2011, The College Board data for the graduating high school classes of 2004, 2006, and 2007 were merged with National Student Clearinghouse (NSC) data, which traces students' postsecondary careers.[10] It identifies all colleges to which the students enroll, whether they transfer to other colleges, and if applicable, when they graduate. Since, these three cohorts were merged in 2011, the 2004 cohort can trace students seven years after high school graduation while students from the 2007 cohort can only be observed for four years after high school graduation.

Finally, I merge these data with the Integrated Postsecondary Education Data System (IPEDS) to generate the final data set, which contains information on the colleges to which students apply and enroll. This information includes type of institution, two-year versus four-year, median SAT score of matriculates, acceptance rate, six-year graduation

---

[9] Throughout, I interchangeably use the terms "Score Send and "application" despite the potential differences.

[10] NSC contains information from over 3,300 colleges, which covers 96 percent of the student population.

rate, freshman retention rate, expenditures per student, faculty per student, and average net tuition.[11]

2.2. Identifying Twins

From this final data set, I identify twins by matching students within the same high school who share the same last name and date of birth. To verify that these linking criteria are adequate, I visually inspected the home addresses, which cannot easily serve as a formal linking criterion due to extensive use of shorthand entry (e.g. Eighth St. versus 8th Street). This results in approximately 30,000 sets of twins. This does not identify all twins because some twins are likely to attend different high schools or one twin may not have used a College Board product. I also check that the twins have distinct social security numbers so that I do not pick up the same person twice and declare twins.

After identifying all possible twins in the data, I keep only the sets of twins in which both students went to a four-year college and both twins sent scores to at least one four-year college.[12] I also eliminate the 2004 students who matriculate three years or

---

[11] Median SAT of enrolled students is approximated by taking the average of A) the sum of the 25th percentile scores on critical reading and math and B) the sum of the 75th percentile scores on critical reading and math. In the few instances that a college reports only average ACT scores, the ACT is converted to SAT scores using the concordance table available here: http://www.act.org/aap/concordance/.

[12] Some students may have enrolled in a four-year college that does not participate in NSC and would consequently be eliminated from the sample.

more after high school graduation, the 2006 students who matriculate after two years, and the 2007 students who matriculate after one year and the corresponding twin. This ensures that both twins have the opportunity to graduate from college in four years, though, many do not. I also exclude some students who are missing critical information, such as gender. This leaves 11,008 sets of twins.

I am unable to identify whether twins are identical or fraternal twins. However, male-female pairs must be fraternal twins and the male-male and female-female comprise of both.

2.3. Descriptive Information

Table 1 presents summary statistics for twins, the differences between sets of twins, and the full sample.[13] The mean SAT score for a student's last SAT attempt is 1118.[14] Within sets of twins, the average difference between SAT scores is 105.4 points. These differences are displayed in Figure 1, which is a scatter plot with one twin's SAT score plotted against the other's. There is a clear linear relationship between the test scores that follows the 45 degree line, and approximately 52 percent of the variability in one twin's SAT score is explained by variability in the other's. The final column in Table 1 shows that overall, the twins are quite similar to the full sample of all four-year college-going students, who have an average SAT of just over 1110.

---

[13] The full sample excludes students for the same reasons that twins are excluded.

[14] SAT is the sum of the scores on the verbal and math sections. In 2006, a writing score was added, increasing the maximum SAT from 1600 to 2400. I do not use the writing score even when available.

The college's selectivity is measured by the institution's median SAT score, and this measure is commonly used as a proxy for selectivity (Dale and Krueger 2002; Long 2008). This is only one imperfect measure of selectivity and alternative measures, including acceptance rate, expenditures per student, graduation rates, persistence rates, student-faculty ratio, and average net tuition are tested. However, all selectivity measures are highly correlated with one another.

Despite the similarity in SAT scores between the twin and non-twin subsamples, there exist sizeable differences in the main outcome measure- four-year graduation rates. 52 percent of sampled twins graduate in four-years, whereas the graduation rate among the non-twin subsample is about 45 percent.[15] The twin and non-twin subsamples also differ somewhat on academic measures like self-reported high school GPA and AP participation. On both of these measures, twins exceed non-twins. Twins also tend to have wealthier parents, with the twin subsample reporting family incomes $6,000 above the non-twin subsample.[16]

---

[15] This variable is whether a student graduates from any college in four years, not just the initially enrolled college. Results are estimated and largely unchanged when using whether graduated in four years from initially enrolled college and consequently not presented.

[16] Income is reported in $10,000 buckets up to $100,000 and then top-coded at over $100,000. For each student, the midpoint of each bucket is used and the top income is set to $120,000.

In addition to the aforementioned variables there exist several other family level variables, which are used in future analyses. Family characteristics that vary between, but not within, sets of twins include ethnicity, state residence, native language, citizenship, father's and mother's educations, county unemployment rate, county educational attainment (percent of population with at least a bachelor's degree), and the student's high school's percent free and reduced price lunch. If students are missing any of these variables, the value is coded as zero and an indicator for missing value is created.

2.4. National Education Longitudinal Study of 1988

The National Education Longitudinal Study of 1988 (NELS) follows students from eighth grade in 1988 as they transition into college in and into the labor force. I observe student demographics and high school academic achievement. I also observe up to three colleges to which they apply, if they are accepted, and whether they enroll. I also observe whether they receive a bachelor's degree by 2000.

From the initial sample I only use students with complete information who enroll in a four-year college that reports their median SAT in IPEDS.[17] The final sample consists of 1,575 students, which are used in a robustness test to replicate the Long (2008) analysis.

3. Conceptual Framework and Empirical Strategy

3.1. Conceptual Framework

[17] Mark Long generously provided code to replicate the analysis. The exact numbers and variables could not be reproduced but come reasonably close.

Let g_{1i}^* and g_{2i}^* equal the propensity of the first and second twin, respectively, in the *i*th pair of twins to graduate from a four-year college. Let X_i represent observable characteristics that vary across sets of twins. This includes parental income, ethnicity, state residence, native language, citizenship, father's and mother's education, and cohort. There are also unobservable characteristics of sets of twins, denoted μ_i.

Let Z_{1i} and Z_{2i} represent variables that vary between and within sets of twins, such as median SAT of college enrolled, student SAT score, number of AP tests taken, number of SAT2's taken, number of Score Sends, and gender. Let P_{1i} and P_{2i} represent application portfolio characteristics that also vary within and between twin sets, such as the number and selectivity of the schools to which a student applies.

College outcomes may also be influenced by student characteristics that are unobservable to the researcher, which are denoted by ε_{1i} and ε_{2i}. This leads to the following equations:

(1) $$g_{1i}^* = \alpha X_i + \mu_i + \beta Z_{1i} + \gamma P_{1i} + \varepsilon_{1i}$$

and

(2) $$g_{2i}^* = \alpha X_i + \mu_i + \beta Z_{2i} + \gamma P_{2i} + \varepsilon_{2i}$$

One advantage to having data on twins, is that the difference between equations (1) and (2) is as follows:

(3) $$g_{1i}^* - g_{2i}^* = \beta(Z_{1i} - Z_{2i}) + \gamma(P_{1i} - P_{2i}) + \varepsilon_{1i} - \varepsilon_{2i}$$

Conveniently, the unobservable family characteristics μ_i has been eliminated from the equation.

3.2. Empirical Strategy

This study strives to obtain an unbiased estimate of β, which in this context represents the coefficient on college selectivity, as measured by median SAT of college enrolled. I observe and set $g_{1i} = 1$ (or $g_{2i} = 1$) if the student graduates in four years, and $g_{1i} = 0$ (or $g_{2i} = 0$) otherwise.

There will be two main estimation strategies: OLS and twin fixed effects. In each strategy, I incorporate specifications that exclude and include application portfolio controls. The four methods allow for a comparison of previous methodologies and take the research a step further by combining methodologies.

3.2.1. OLS

I estimate β by pooling equations (1) and (2) and running OLS. Standard errors are clustered at the twin level.[18]

Within OLS, I use two specifications. The first excludes controls for application portfolio characteristics, P_{1i} and P_{2i}, and the second includes them. Portfolio characteristic controls include number of Score Sends, and the minimum, mean, and maximum of the median SAT score of enrolled students for the colleges in the application portfolio.[19]

---

[18] Multi-way clustering at the twin level and college of matriculation level has no effect on results and consequently withheld.

[19] I also test for different forms of portfolio controls including just the mean of the median SAT score and number of applications (both the linear count and non-linear count using dummies, as in Dale and Krueger 2002).

Including these portfolio controls is in the spirit of Dale and Krueger (2002). In their seminal paper, which estimates the impact of college selectivity on wages, they control for the set of schools to which students apply and are accepted. Their main specification includes portfolio fixed effects by matching students with similar or exact portfolios and acceptance outcomes. They also use a "self-revelation model," which controls for the number of applications and mean selectivity of those applications. The two approaches yield similar results in their paper. This paper uses the latter approach for ease and precision but a variant of the former approach is tested and presented as a robustness check and results are consistent across the two approaches.

3.2.2. Twin Fixed Effects

By utilizing the sets of twins, I difference out the unobservable family effect in equation (3). Formally, I use twin fixed effects.[20] Therefore, identification comes from variation within a set of twins. Table 1's column of differences within twins suggest that there is substantial variation to exploit. Similarly, Figure 2 plots the relationship between twins' differences in the median SAT scores of enrolled college (selectivity) and differences in graduation rates. The upward sloping line is suggestive that the twin who enrolls in the more selective school is more likely to graduate. Of course, this does not control for other important characteristics that the model includes so as to get the true causal relationship.

[20] When the group size equals two, first differences and fixed effects models are equivalent.

Again, I use two specifications: excluding and including application portfolio controls. And relative to OLS, observable family characteristics (e.g. income, ethnicity, state, cohort, etc.) must be excluded.

3.2.3. Identification

OLS

When running OLS, I use the variation in enrollment both between and within twins to estimate the effect of college selectivity on four-year graduation rates. The sample includes approximately 1,000 different colleges of varying selectivity to exploit. Even within twins, only 25 percent of the pairs attend the same college, so there is variation within the family as well. This variation is conditional on student, parent, and local characteristics. Assuming equations (1) and (2) are correctly specified, one must assume that $\mu_i = 0$ or that $Corr(\mu_i, Z_{1i}) = Corr(\mu_i, Z_{2i}) = 0$ to get an unbiased estimate of β. In other words, is there something, such as an unobserved parental characteristic or student aspiration, which influences both a student's enrollment decision and the likelihood of graduating? Correctly specifying equations (1) and (2) means that there can be no student level or twin level unobservables that are correlated with the median SAT score of the enrolled college to obtain causal estimates. If there do exist unobservables that are correlated with median SAT, there is likely a positive relationship. For example, motivation or parental support may push students into a more selective college and at the same time push students to graduate, resulting in a positive bias. These valid concerns motivate other specifications.

Application Portfolio Controls

Next, I include controls for application portfolio characteristics, P_{1i} and P_{2i}. Dale and Krueger (2002) argue that by including application portfolio controls (or fixed effects), the econometrician can compare two students with the same motivation or aspiration. Identification comes from variation within students who have the same portfolio. The hope is that conditional on the same portfolio and student characteristics, there are no unobservables that affect the probability of graduating and that also correlate with the enrollment decision. Thus, exploiting this within portfolio variation may reduce the aforementioned positive selection bias (e.g. motivation and aspiration). However, there are two potential threats to this strategy. First, admission outcomes dictate where a student enrolls and may be correlated with whether a student graduates. Second, the unobserved decision rule that students use to matriculate may be correlated with graduation outcomes. These are discussed in turn.

As for admission outcomes, unlike Dale and Krueger (2002) and Long (2008), I do not observe where students are admitted, only where they apply. Dale and Krueger (2002) find that only controlling for the applications (the "self-revelation" model) yields results consistent with those that also control for admission decisions. Despite the consistency across specifications, they argue that including controls for where students are accepted controls for student ability that is unobservable to the econometrician but is observable to the admissions committees: a potential source of bias. For example, students who write strong admissions essays may be more likely to graduate than observationally identical students, simply because a strong essay is an indicator of talent or motivation that is not easily quantifiable. To mitigate this issue, I include as many controls as possible including student, family, and local characteristics.

In addition, I test the sensitivity of results to excluding key criteria (e.g. student SAT and high school GPA), conditional on the application portfolio, to see whether the results change relative to including a full set of controls. As seen in the results, selectivity estimates do not substantially increase, which suggests that the application portfolio does a reasonable job at controlling for unobservable student characteristics.

Next, I return to the data in Long (2008), the National Education Longitudinal Study of 1988, to investigate how sensitive results are to only having applications and not admissions outcomes. I attempt to replicate his result, which is a replication of Dale and Krueger (2002), but using bachelor degree attainment as the outcome variable and students are matched on applications and admission decisions. I then estimate the same specification but only match students on application portfolios. As seen in the robustness results, including admission outcomes does not statistically alter the results, thus providing evidence that using only application controls does not substantially bias results.

As a final argument that lack of admission outcomes do not threaten identification, note that there may exist randomness in admissions decisions unrelated to unobservable student ability. For example, two students identical on all measures related to graduation rate probability, may write essays of identical quality that differ in their appeal to an admissions officer. Hence, two identical students may have different choice sets by virtue of an essay topic. As another example, two students may be identical but live in different states. If they apply to the same schools they may have different admission outcomes based on preferences of the schools or cohorts sizes within states. And as a final example, one student may get a 1200 on her SATs and another student may get an 1190, which may produce differential choice sets.

As for matriculation decisions, it is difficult to fully rule out that there is nothing unobservable about the student who enrolls in the more selective college that is correlated with likelihood of graduating. But there are certainly students' horizontal preferences that are not related to quality. That is, students have preferences for college attributes that are unrelated to academic quality, which ultimately drives their enrollment decision (Jacob, McCall, & Stange, 2013). These attributes may include sports teams (Pope and Pope, 2009), regions of the country, programs of study, dormitory quality, whether a friend attends, etc., none of which have to be associated with selectivity. Hence two observationally identical students end up at two schools of differing selectivity, but not because one is more motivated than the other.

Overall, once the application is controlled for, it is assumed that student ability and motivation are properly measured. Hence, identification comes from randomness in student and college preferences that are unrelated to unobservable student characteristics that affect the probability of graduating.[21]

Twin Fixed Effects

Including twin fixed effects eliminates both observed and unobserved family characteristics that are fixed within a family that may influence enrollment and graduation decisions. This relaxes the assumptions that $\mu_i = 0$ and the assumption that μ_i must be uncorrelated with Z_{1i} and Z_{2i}. Hence, identification comes from variation in school enrollment choices within sets of twins. Thus, exploiting this within twin

---

[21] If the assumptions are not met it is most likely that the unobservable attributes positively biases the estimates on return to selectivity.

variation may reduce the aforementioned positive selection bias (e.g. parental motivation, environmental factors, and family wealth). And when controlling for the application portfolio, the variation is within a family and within a portfolio.

3.2.4. Critiques of Twin Fixed Effects

There exist numerous critiques to using within twin variation as an identification strategy. The first critique is whether twins are representative of the overall population and consequently whether any estimates have external validity. The aforementioned descriptive statistics suggest twins are slightly more likely to graduate from college than the overall population, but beyond that, they are quite similar. Also, these twins were typically born between 1986 and 1989, before the proliferation of fertility treatment, particularly among wealthier families.

The second critique is if these twins are so similar, why do they make such different enrollment decisions? Research across several disciplines discusses the genetic and environmental differences among twins (e.g. Jencks, 1980; Kamin and Goldberger, 2001; Manski 2011). Econometrically, this amounts to the exogeneity of the within-twin variation or more specifically, the existence of unobservable student-level characteristics that influence college choice, and independently, the probability of four-year graduation rates. This paper assumes that conditional on observables, the enrollment decisions of twin pairs are exogenous. Estimates are likely positively biased if this assumption does not hold and therefore it is important to mitigate the issue by including a rich set of student-level controls. If the set of controls does not account for unobserved ability differences, both Neumark (1999) and Bound and Solon (1999) point out that using twin fixed effects can exacerbate bias relative to OLS estimates. While it is difficult to conclusively support the exogeneity assumption, in corroboration of it, I find the twin with a lower SAT than the other twin enrolls in a more selective college 46 percent of the time. This is consistent with twins having different preferences for types of colleges or programs that are unrelated to student ability or college quality.

4. Results

Table 2 presents the main set of results. Column (1) presents results when using the full sample of SAT test takers and regressing whether the student graduates on the median SAT of college enrolled with no controls. The coefficient of 0.121 and very small standard errors imply that attending a college where the median student SAT is 100 points above the alternative college is associated with a 12.1 percentage point increase in graduation probability.

Column (2) controls for student and family varying characteristics and the coefficient on median SAT of college enrolled drops to 0.054. Also, the estimated coefficient is approximately four times the size of the estimated coefficient on a student's SAT. Column (3) includes controls for the type of application portfolio, and results in a very similar estimate of 0.051.[22] This suggests that, after controlling for student

[22] Excluding controls for the minimum and maximum SAT in the portfolio does not change results. Neither does using dummies for the number of applications.

demographics and measures of academic achievement, the composition of the application portfolio has very little explanatory power.[23]

The next three columns run the same regressions but only for the sets of twins. Coefficient estimates in columns (5) and (6) are over 5 percentage points and statistically significant. Similar to the full sample estimate, portfolio characteristics controls do not change the estimates in a meaningful way.

Columns (7) through (9) are results from the twin fixed effects models. The coefficient estimates here are 0.087, 0.052 and 0.048 when moving from fewer controls to more controls. These estimates are comparable to the OLS estimates.

Overall, these results suggest that, when choosing between two colleges where the typical students differ by 100 SAT points, the student will enjoy a five percentage point increase in graduation probability by attending the college with the higher median SAT score. Moreover, the results also suggest that many of the major selection issues that researchers discuss, such as family unobservables and unobservable desire to attend and graduate from college, do not have a strong impact on the estimates.

4.1. Robustness Checks

4.1.1. Portfolio Fixed Effects

The first robustness check identifies whether the previous controls for application portfolios are too restrictive. Hence, I follow the methodology of Dale and Krueger

[23] Controlling for more higher order terms of student's SAT does not impact these results or future results.

(2002) by creating dummy variables for students with similar portfolios. I then re-estimate the models but only include students that have at least one other student who has a similar portfolio.[24]

To construct similar portfolios I use the median SAT of enrolled students for every school that received a Score Send.[25] Schools are then placed into 25 SAT point buckets. For example, there is a bucket for schools with median SATs between 1101 and1125 and a bucket for schools with median SATs between 1126 and 1150. This generates 34 buckets. For each student, I construct the count of Score Sends in each of the 34 buckets. A student's portfolio is described by the 34 element vector. I then create a set of dummy variables for each unique vector. I only include students who have the same portfolio as at least one other student.[26]

---

[24] I present results that include students who only apply to one school while Dale and Krueger (2002) exclude these students. The results are not sensitive to the exclusion of them.

[25] This excludes two-year institutions and the few non-academic institutions (e.g. financial aid institutions). Some schools only report the average ACT score but I convert them with the aforementioned ACT/SAT concordance table.

[26] Including students who had a portfolio that no one else applied to does not affect estimates because there is no within portfolio variation when including portfolio fixed effects.

Results are presented in Table 3. The first three columns use the full sample of students with matched portfolios. The first column yields an estimate of 0.120 and once student and family characteristics are controlled for the estimate drops to 0.052 and once portfolio fixed effects are included, the estimate drops slightly to 0.049. Using the sample of twins and OLS gets very similar results. Column (8) presents the coefficient estimate of 0.054 when using twin fixed effects but no portfolio fixed effects. The last column, which has both fixed effects, actually presents an increase in the estimate to 0.063.

Including portfolio fixed effects is much less restrictive but comes at the cost of a smaller sample and less precision. However, estimates across the table are very near five percentage points and statistically different than zero.

4.1.2. Alternative Measures of Selectivity

Median SAT of enrolled students is an imperfect measure of college selectivity. Consequently, I re-run the regressions with five different measures of selectivity: admission rejection rate, six-year graduation rate, freshman retention rate, expenditures per student, faculty to student ratio, and average net tuition. These selectivity measures are also imperfect but collectively, if they yield similar results to the previous table, then the estimated effect is strongly supported. For comparison's sake, I normalize all variables to mean zero with a unit standard deviation and also redisplay results for the normalized median SAT of the college (one standard deviation is approximately 100 SAT points).

Table 4 presents results from the four main regressions and only use the twin sample. A clear pattern appears when scanning the rows: all estimates are statistically significant and directionally consistent with the median SAT result.

Also, the magnitudes of the coefficients in a given column are not far from one another, but there are some statistical differences. The differences may exist because they are noisy measures of selectivity. To address this, I use factor analysis to produce a college selectivity index. Black and Smith (2006) suggest this method to address the highly correlated but imperfect measures of selectivity that can lead to biased results. The last row demonstrates that a one standard deviation increase in selectivity is associated with approximately a 7 percentage point increase in the probability of graduating.

There is also a clear pattern when scanning across the columns: including twin fixed effects and portfolio controls reduces the coefficient estimates' magnitudes. That is, the last column's estimates are all smaller in magnitude than the first column's estimates. Together with Table 2, there is evidence that OLS results are biased upwards, but not by much.

Finally, it is well established that attenuation bias occurs when there is measurement error but it can be even greater with fixed effects (Griliches, 1979). This would imply that the fixed effects results could be smaller than the OLS estimates, which is not true in this table. If in fact, attenuation bias is driving down fixed effects estimates, then it would necessarily be true that OLS estimates are biased upwards since we have stable estimates. It is impossible to disentangle these forces but even if there is measurement error in each measure of selectivity, there is a positive estimate on attending a more selective college.

4.1.3. Sensitivity to Admission Predictors

I now test whether omission of variables that admission committees are likely to value changes the estimates. In doing so, I am testing the magnitude of the potential bias from not observing admission outcomes, conditional on the application portfolio.

Specification (1) of Table 5 presents the main result from Table 2, which includes a full set of controls and twin fixed effects. The next specification excludes student SAT, followed by excluding high school GPA, and then both. Excluding these critical measures of admissions committees does inflate the estimate on Median SAT of Enrolled College, but not by much and not statistically so. Column (5) excludes all academic measures available and is close to being statistically larger than the specification with a full set of controls.

These results indicate that controlling for the application portfolio gives a good indication as to the student's ability, at least relative to traditional measures of ability, such as SAT and GPA. It may still be the case that admissions committees observe something else that the econometrician does not (e.g. admission essay or letters of recommendation), which determines where students enroll and may be correlated with whether they graduate. However, SAT and GPA are very strong candidates for being correlated with the quality of the college enrolled and potential sources of bias.

4.1.4. Excluding Admissions Outcomes with NELS

Table 6 presents results using NELS. And the first two columns show Long (2008) results. He finds a marginal effect on attending a more selective college, as measured by a normalized median SAT, which is not statistically different than the Dale and Krueger (2002) method that uses application and admission fixed effects.

The next two columns present my attempt at replicating his results.[27] The results are larger than Long (2008) but are not statistically different than one another. Moreover, they both show a smaller fixed effect estimate than OLS, but no statistical differences. The last column runs the analysis but only matches students on applications. The two fixed effects models are not statistically different from one another and neither of them are statistically different than OLS. This suggests that, in this context, controlling for the application alone does not change estimates relative to controlling for application and admission outcomes. This may in part be due to the imprecise estimates that include fixed effects. But it would be troubling if the coefficient in the last column was much larger (positively biased) than the Dale and Krueger method.

5. Extensions

5.1. Nonlinear Effects

The previous results assumed a linear relationship between median college SAT and graduation probability, but the true effect may be nonlinear. To capture the potential nonlinearities, I create several dummy variables for the median SAT of the school enrolled: less than 1000, between 1000 and 1099, between 1100 and 1199, between 1200

[27] Long (2008) and this analysis contain different students likely due to slightly different selection criteria. There are also a few differences in available variables.

and 1299, and greater than or equal to 1300.[28] Table 7 presents results of the four main models that use the twin sample. The omitted category includes schools where the median SAT is less than 1000-the least selective schools.

The first column is the result of running OLS. Students who enroll in the most selective schools are about 20 percentage points more likely to graduate than those who enroll in the least selective schools. There is a substantial 10 percentage point increase in the magnitude of the coefficient when going from schools with a median SAT between 1000 and 1099 to those with a median SAT between 1100 and 1199. All other coefficients change by approximately 3 to 4 percentage points across categories.

When looking at the next few columns, a few results stand out. First, there always appears to be a noticeable increase in the graduation rate when going from schools between 1000-1099 and 1100-1199, even with more controls. Second, the differences between schools with median SATs less than 1000 and 1000-1099 are small (and statistically insignificant). Third, there are no statistical differences between the most selective schools and schools with median SATs between 1200 and 1299.

Overall, the full model with twin fixed effects and application portfolio controls implies that there are nonlinearities. The nonlinearities suggest that the largest gains in graduation probability occur when choosing moderately selective colleges over less selective colleges, rather than highly selective colleges over moderately selective colleges.

---

[28] This represents 175, 268, 199, 85, and 70 distinct colleges, respectively. In addition, the college six-year graduation rates are 0.42, 0.52, 0.64, 0.74, and 0.88, respectively.

5.2. Heterogeneous Effects

This subsection uses subsamples to test whether there are heterogeneous effects on the outcome by gender, race/ethnicity, parental income, and high school urbanicity.[29] The four main specifications are used.

5.2.1. Gender

I first split the sample by twins' genders: male only, female only, and mixed gender sets. This exercise offers two pieces of information. First, it sheds light on who is driving the previous results. Second, same sex twins are comprised of identical and non-identical twins (approximately one-third of same sex twins are identical). But identical twins are, as the name implies, genetically identical. Berman, Rosensweig, and Taubman (1996) argue that identical twins not only eliminate family fixed effects, but also genetic fixed effects that may bias results. Hence, if the coefficient estimates substantially differ between the mixed gender and same gender twins, there is suggestive evidence that genetic differences play an important role in the estimation.

In fact, the average difference in SAT among male-male twins is 95 points and 92 points for female twins. This suggests that there are more similarities among same sex twins, but there are still substantial differences, perhaps driven mostly by the fraternal same sex twins.

---

[29] Some students have missing data and are excluded so subsample counts do not necessarily sum to the full twin sample.

Table 8 presents results of the estimation by gender subsample. Using male only twins, the full model's estimate is 0.061 and this does not differ much from the basic OLS estimate. On the other hand, female only twins have an estimate of only 0.032, which is also substantially lower than the OLS estimate. Finally, the mixed gender twins have an estimate of 0.056.

Combined, this is evidence that college selectivity has a greater impact on graduation probability for males, compared to females. Also, the unobservable family fixed effect is stronger for female twins. Moreover, mixed gender twins estimates fall in between the same sex estimates. Though it is impossible to disentangle the mixed gender effect from the non-identical twin effect, this is supportive evidence that the non-identical effect is not driving results.

5.2.2. Race, Income, Urbanicity

The next few results parse the data by race, income, and high school urbanicity. It has been shown that under-represented minorities, low-income students, and students in rural high schools have relatively undesirable college going processes and completion rates (Bowen, Chingos, and McPherson, 2009; Dillon and Smith, 2009; Smith, Pender, and Howell, 2013; Hoxby and Avery, 2013). The first subset of results uses only White students and shows a coefficient of 0.046 in the full model. The analogous coefficient in the Black/Hispanic subsample is 0.030, but is less precisely estimated and not statistically different from the White students' coefficient.

When comparing results across income groups, there are larger coefficients for wealthier students compared to less wealthy students. Finally, when stratified by the urbanicity of the student's high school, all students' impact estimates are positive and statistically significant in the full model, but the suburban high school estimate is at least 60 percent greater than all other coefficients.

5.3. Longer Term Graduation Rates

Thus far, I use whether or not a student graduates from college in four years as the outcome variable. But only 38 percent of students at four-year institutions graduate in four years, whereas 54 percent and 58 percent graduate within five and six years, respectively (NCES 2011).[30] Hence, the effects of college selectivity may have differential effects when examining longer term outcomes.

Table 9 compares results for varying degree attainment lengths.[31] In the OLS specification, there are no statistical differences between the coefficient estimates. However, with both twin fixed effects and application portfolio controls, the estimates are larger for five year graduation rates (5.8 percentage points) and still even larger for six-year graduation rates (7.8 percentage points). In addition, Appendix 1 shows that the non-linear estimates of five- and six-year graduation rates are larger in magnitude than the four-year graduation rates (with a substantial 21 percentage point increase in the probability of longer term graduation when enrolling in a college with median SAT

---

[30] The 38 percent four year graduation rate is lower than the graduation rates in this analysis because I use relatively motivated students who enroll soon after high school and take the SATs.

[31] The samples differ because longer term outcomes can only be considered for earlier cohorts.

greater than 1100, as opposed to a less selective college). This evidence suggests that the previous results may be underestimating the true effect of enrolling in a relatively selective institution.[32]

5.4. Undermatch/Overmatch

Finally, I compare whether the selectivity of the institution, relative to the student's own SAT has an effect on graduation probabilities. This is in the spirit of the undermatch and overmatch literatures.[33] Formally, I subtract the median SAT at the college to which the student enrolls from the student's SAT. I then create several indicators for whether there is a substantial difference between the two (e.g. more than 100 SAT points).

Table 10 displays the results of this exercise. The first three rows have different measures of undermatch but have relatively consistent results: undermatching can negatively affect the probability of graduating. However, there are not big differences in

---

[32] Appendices 1 and 2 shows that some of the non-linear and linear estimates, respectively, are driven by differences in cohorts since the 2004 cohort has smaller coefficients (on the four-year graduation estimates) than when using all cohorts.

[33] Undermatch first appeared in Roderick et al. (2006) and *Crossing the Finish Line* (2009). Overmatch is typically in the context of affirmative action. For a complete review of both literatures, see Smith and Pender (2012). Strictly speaking, this is not undermatch and overmatch because it does not account for where students could enroll. But relative SAT scores is a good proxy.

the effects across the different magnitudes of undermatch. That is, undermatching by more than 50 points is associated with a 5.6 percentage point reduced probability of graduating but is not estimated to have substantially different effects than undermatching by more than 200 points.

Relative to undermatching estimates, the overmatching estimates are smaller in magnitude and the opposite sign. There is suggestive evidence of positive benefits to overmatching, but it is not strong.

6. Conclusion

This paper has presented several results that deserve further consideration. First, the results show that enrolling in a relatively selective college increases the probability of graduating. I estimate that enrolling in a school with a median SAT score 100 points greater than an alternative school increases the probability of graduating in four years by approximately five percentage points. Long (2008) finds convincing evidence of a similar effect using instrumental variables, propensity score matching, and the Dale and Krueger (2002) method.

Second, the results are quite stable across several methodologies, including twin fixed effects and controlling for application portfolios. The stability is reassuring for researchers who do not have access to such a large set of controls but are interested in unbiased estimates between college selectivity and a host of other outcomes. The stability slightly differs from Long (2008), who identifies estimate sensitivity across models. Perhaps one contributing factor to the difference between our results is that the cohorts are at least twelve years apart. An alternative explanation is that observation of admission outcomes reduces the bias, however, I show evidence that missing key admission criteria does not substantially change results.

Third, related research is often concerned with wages, which I do not observe. Several papers find mixed evidence that there is a small but positive effect of college quality on wages, at least for some subset of students (Dale and Krueger 1994; Behrman, Rosensweig, and Taubman 1996; Long 2008). These small and often statistically insignificant estimates are consistent with the above results. Let's take the median wages of a person with an associate's degree ($42,000) and bachelor's degree ($55,700).[34] If there is a five percentage point increase in probability of graduating by enrolling in a more selective school, then there is a 1.4 percent increase in expected wages. So if an increased probability of graduation is the only mechanism by which wages are higher for those who attend more selective colleges, these results are consistent with previous work since there are small effects on wages that would be difficult to precisely estimate.

Fourth, despite the small back of the envelope calculation on wages, there may be other benefits to enrolling in a more selective college. Those who do attend college are more likely to have a healthier lifestyle, employment and insurance, vote, and volunteer (College Board 2012). There also exist differences between individuals who graduate college and those who attend some college (i.e. leaving before graduating) or get an Associate's degree. There is also the consumption value of college, such as students enjoying sports, activities, friendships, and other non-academic aspects of college, which

[34] For full-time workers over the age of 25 in 2008 and it excludes those who get more advanced degrees (Baum, Ma, and Payea 2010).

has been documented to be an important decision for students (Jacob, McCall, & Stange 2013).

Lastly, it should also be noted that five percentage points is quite large-- an approximately 10 percent increase in the probability of graduating in four years and five to six times larger than the estimated effect of student SAT. There are numerous implications for policymakers, depending on their goals. If policymakers want to achieve educational attainment equity on some measure (e.g. race or income), then some responsible party (e.g. parent, teacher, counselor, colleges) should help get students into more selective colleges. Alternatively, policymakers can invest in the relatively low-completion rate colleges, which is where minorities and low-income students disproportionately enroll. Even if equity is not the goal, investment in these low-completion rate colleges may help students graduate and in the labor market.

References

Alon, S. and M. Tienda (2005), "Assessing the 'Mismatch' Hypothesis: Differentials in College Graduation Rates by Institutional Selectivity." *Sociology of Education*. 78(4): 294- 315.

Ashenfelter, O. and A. Krueger (1994). “Estimates on the Return to Schooling From a New Sample of Twins,” *American Economic Review*. 84: 1157-1173.

Hoxby, C. and C. Avery (2013). “The Missing One-Offs: The Hidden Supply of High-Achieving, Low Income Students,” NBER Working Paper #18586.

Baum, S. J. Ma, and K. Payea (2010). “Education Pays 2010: The Benefits for Higher Education for Individuals and Society,” College Board – Advocacy & Policy Center.

Behrman, J., M. Rosensweig, and P. Taubman (1996). “College Choice and Wages: Estimates Using Data on Female Twins,” *The Review of Economics and Statistics*. 78(4): 672-685.

Black, D. A. and J. A. Smith (2004). “How Robust is the Evidence on the Effects of College Quality? Evidence From Matching.” *Journal of Econometrics*. 121. 99 – 124.

Black, D. A. and J. 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 Solon, G. (1999). “Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling,” *Economics of Education Review*, 18: 169-182.

Bowen, W. and D. Bok (1998). *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*. Princeton University Press, Princeton, NJ.

Bowen, W.G., M.M. Chingos, and M.S. McPherson (2009). *Crossing the finish line*. Princeton: Princeton University Press.

Brand, J., and C. Halaby (2006). "Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement" *Social Science Research*, 35(3): 749-770.

Brewer, D., E. Eide, and R. Ehrenberg (1999). "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources*, 34(1): 104-23.

Card, D. and A. Krueger (2005). ”Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants? Evidence From California and Texas,” *Industrial and Labor Relations Review*, 56: 416-434.

Cohodes, S. & J. Goodman. (2012). “First Degree Earns: The Impact of College Quality on College Completion Rates.” Faculty Research Working Paper Series. Kennedy School of Government, Harvard University.

College Board (2012). “Five Ways Ed Pays.” Accessed online July 25, 2012: http://advocacy.collegeboard.org/five-ways-ed-pays/home.

Dale, S. and A. 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): 1491-1527.

Dale, S. and A. Krueger. (2011). "Estimating the Return to College Selectivity over the Career Using Administrative Earning Data," Working paper.

Daniel, K., D. Black, and J. Smith (1997). "College Quality and the Wages of Young Men." *University of Western Ontario*, Department of Economics, Research Report Number 9707.

Dillon, E. and J. Smith (2009). "The Determinants of Mismatch Between Students and Colleges," (Unpublished working paper). Ann Arbor, MI: University of Michigan.

Eide, E., D. Brewer, and R. Ehrenberg (1998). "Does it Pay to Attend an Elite Private College? Evidence on the Effects of Undergraduate College Quality on Graduate School Attendance." *Economics of Education Review*. 17(4): 371-376.

Griliches, Z. (1979). "Sibling Models and Data in Economics: Beginnings of a Survey." *The Journal of Political Economy*, 87 (5): 37-64.

Hoekstra, M. (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.

Horn, L. and Carroll, C.D. (2006). “Placing College Graduation Rates in Context: How 4-year College Graduation Rates Vary With Selectivity and the Size of Low-Income Enrollment.” *Postsecondary Education Descriptive Analysis Report*. National Center for Education Statistics, U.S. Department of Education.

Hoxby, C. (1998). "The Return to Attending a More Selective College: 1960 to the Present." Mimeo, *Harvard University*.

Jacob, B., B. McCall, and K. Stange (2013). “College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?” NBER Working Paper #18745

James, E., N. Alsalam, J. Conaty, and D. To (1989). "College Quality and Future Earnings: Where Should You Send Your Child to College?" *The American Economic Review*. 79(2): 247 – 252.

Kane, T (1998). "Racial and Ethnic Preferences in College Admissions." *The Black-White Test Score Gap*, Edited by Jencks, Christopher and Meredith Phillips, Brookings Institution Press, Washington, D.C., pp. 431-456.

Lindahl, L., H. Regner. (2005). “College Choice and Subsequent Earnings: Results Using Swedish Sibling Data,” *Scandinavian Journal of Economics* 107(3): 437-457.

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

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

Manski, C. (2011). “Genes, Eyeglasses, and Social Policy,” *Journal of Economic Perspectives*, 25(4): 83-94.

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

NCES (2011). *Digest of Education Statistics*. Table 345. Accessed online 7/23/12: http://nces.ed.gov/programs/digest/d11/tables/dt11_345.asp.

Neumark, D. (1999). “Biases in Twin Estimates of the Return to Schooling,” *Economics of Education Review*, 18(2): 143-148

Pallais, A. (2012). "Small Differences That Matter: Making Mistakes in Applying to College," Working Paper.

Pope, D. and J. Pope (2009). "The Impact of College Sports Success on the Quantity and Quality of Student Applications," *Southern Economic Journal*, 75(3): 750-780.

Roderick, M., Nagaoka, J., Allensworth, E., Coca, V., Correa, M., Stoker, G. 2006. "From High School to the Future: A First Look at Chicago Public School Graduates' College Enrollment, College Preparation, and Graduation from Four-Year Colleges." Consortium on Chicago School Research, Chicago, March.

Rouse, C. (1999). "Further Estimates of the Economic Return to Schooling From a New Sample of Twins," *Economics of Education Review*. 18(2): 149-157.

Saavedra, J. (2008). "The Returns to College Quality: A Regression Discontinuity Analysis," Working paper.

Smith, J.I. & Pender, M. 2012. Student-College Academic Match – Literature and Landscape Review. College Board Advocacy and Policy Center. Working Paper.

Smith, J., M. Pender, and J. Howell (2013). "The Full Extent of Academic Undermatch," *Economics of Education Review*, 32: 247-261

Zhang, L. (2005). "Advance to Graduate Education: The Effect of College Quality and Undergraduate Majors." *Review of Higher Education*. 28(3): 313-338.

Table 1: Summary Statistics

Variable	Twins Levels	Twins Differences	Full Sample Levels
Student SAT (100s)	11.180	1.054	11.105
	(1.875)	(0.923)	(1.879)
Median SAT of College Enrolled (100s)	11.454	0.602	11.323
	(1.279)	(0.826)	(1.285)
Graduated College in Four Years	0.518	0.301	0.449
	(0.500)	(0.459)	(0.497)
Number of AP Tests Taken	2.096	0.954	1.840
	(2.523)	(1.418)	(2.375)
Number of SAT2's Taken	0.882	0.361	0.835
	(1.442)	(0.868)	(1.415)
Number of Score Sends	6.034	1.660	5.951
	(3.327)	(2.000)	(3.402)
High School GPA[1]	3.611	0.336	3.526
	(0.521)	(0.374)	(0.542)
Parents' Income[2]	76,012	--	69,773
	(38,680)	--	(39,049)
Observations	22,016	22,016	2,029,483
Sets of Twins	11,008	11,008	--

Notes: Standard deviations in parentheses. Population includes SAT test takers that went to a four-year institution in the 2004, 2006, and 2007 graduating high school cohorts. Twins are identified in the population by date of birth, last name, high school, and street address.

1. Only 10,142 sets of twins and 1,872,468 of the full sample report GPA.

2. Only 6,842 sets of twins and 1,280,908 of the full sample report parents' income.

Table 2: Effect of College Selectivity on Graduation
Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise

	Full Sample			Twins					
	OLS			OLS			Twin Fixed Effects		
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Median SAT of College Enrolled (100s)	0.121***	0.054***	0.050***	0.118***	0.058***	0.052***	0.087***	0.052***	0.048***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.004)	(0.007)	(0.006)	(0.006)	(0.007)
Student SAT (100s)	--	0.014***	0.014***	--	0.009***	0.009***	--	0.010**	0.009*
	--	(0.000)	(0.000)	--	(0.003)	(0.003)	--	(0.005)	(0.005)
Student Achievement Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Family Characteristic Controls	No	Yes	Yes	No	Yes	Yes	No	No	No
Application Portfolio Controls	No	No	Yes	No	No	Yes	No	No	Yes
Observations	2,029,483	2,029,483	2,029,483	22,016	22,016	22,016	22,016	22,016	22,016
R-squared	0.098	0.175	0.176	0.091	0.173	0.175	0.026	0.056	0.057

Notes: Robust standard errors are in parentheses. Twins results cluster standard errors at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. Student achievement controls include SAT score, number of SAT2's taken, and dummies for number of AP tests taken and GPA. Family characteristics include income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, and high school percent free and reduced price lunch, and male and year dummies. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Table 3: Effect of College Selectivity on Graduation - Portfolio Fixed Effects

Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise

Students With Macthed Portfolios[1]

	Full Sample			Twins					
Variable	OLS			OLS			Twin Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Median SAT of College Enrolled (100s)	0.120***	0.052***	0.049***	0.113***	0.053***	0.051***	0.087***	0.054***	0.063***
	(0.000)	(0.000)	(0.000)	(0.005)	(0.007)	(0.010)	(0.005)	(0.013)	(0.020)
Student SAT (100s)	--	0.015***	0.015***	--	0.009*	0.007	--	0.013	0.015
	--	(0.000)	(0.000)	--	(0.004)	(0.006)	--	(0.008)	(0.010)
Student Achievement Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Family Characteristic Controls	No	Yes	Yes	No	Yes	Yes	No	No	No
Portfolio Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes
Observations	1,530,932	1,530,932	1,530,932	7,206	7,206	7,206	7,206	7,206	7,206
R-squared	0.087	0.167	0.253	0.065	0.171	0.508	0.707	0.748	0.857

Notes: Robust standard errors are in parentheses. Twins results cluster standard errors at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. Student achievement controls include score, number of SAT2's taken, and dummies for number of AP tests taken and GPA. Family characteristics include income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, and high school percent free and reduced price lunch, and male and year dummies.

1. Students portfolios are described by the number of Score Sends to each selectivity bucket. Selectivity buckets are 25 SAT point bands (e.g. 1101-1125 and 1126-1150, etc.). A macthed portfolio is when at least two students have the same portfolio.

Table 4: Effect of College Selectivity on Graduation - Alternative Measures of Selectivity

Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise

Each Coefficient Estimate from Separate Regression

All variables are normalized to mean zero with unit standard deviation

Twins Only

College Enrolled Measure of Selectivity:	OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
Median SAT of College Enrolled (100s)[1]	0.074***	0.066***	0.067***	0.062***
	(0.005)	(0.006)	(0.008)	(0.008)
Admission Rejection Rate	0.034***	0.026***	0.037***	0.032***
	(0.004)	(0.004)	(0.007)	(0.007)
Six-Year Graduation Rate	0.095***	0.091***	0.082***	0.078***
	(0.005)	(0.005)	(0.008)	(0.008)
Freshman Retention Rate	0.042***	0.036***	0.044***	0.040***
	(0.004)	(0.005)	(0.008)	(0.008)
Expenditures per Student ($10,000s)	0.056***	0.047***	0.043***	0.039***
	(0.005)	(0.005)	(0.006)	(0.006)
Faculty to Student Ratio	0.047***	0.036***	0.033***	0.029***
	(0.005)	(0.005)	(0.005)	(0.005)
Average Net Tuition	0.068***	0.062***	0.057***	0.056***
	(0.004)	(0.004)	(0.008)	(0.008)
College Selectivity Index[2]	0.094***	0.090***	0.076***	0.074***
	(0.006)	(0.006)	(0.010)	(0.010)

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends. Some colleges do not report selectivity measures so some observations missing.

1. One standard deviation is approximately 100 SAT points.

2. Uses principal component factor analytic method with seven factors to construct a college selectivity index.

Table 5: Sensitivity to Admission Predictors
Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise
Twin Fixed Effect Models

Variable	(1)	(2)	(3)	(4)	(5)
Median SAT of College Enrolled (100s)	0.048***	0.050***	0.054***	0.057***	0.065***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Excluding SAT	No	Yes	No	Yes	Yes
Excluding High School GPA	No	No	Yes	Yes	Yes
Excluding Counts of AP and SAT2	No	No	No	No	Yes
Observations	22,016	22,016	22,016	22,016	22,016
R-squared	0.057	0.056	0.050	0.049	0.044

Notes: Robust standard errors are in parentheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include twin fixed effects, gender dummy, and application portfolio controls, including number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Table 6: Testing Application and Admission Controls

Dependent Variable = 1 if Graduates with Bachelors Degree, 0 Otherwise

Using 1992 High School Cohort from NELS

	Long (2008)		New Estimates		
	OLS	Application and Admission Fixed Effects	OLS	Application and Admission Fixed Effects	Application Fixed Effects
Median SAT of College Enrolled (z-statistic)	0.053	0.025	0.068***	0.053	0.013
	(0.035)	(0.054)	(0.018)	(0.059)	(0.053)
Observations	1505	1505	1575	1575	1575
R-squared	--	--	0.146	0.468	0.358

Notes: Long (2008) results are not estimated but other results use National Education Longitudinal Study data with panel weights. Standard errors in parentheses. *** means significant at 1% level, ** at 5%, and * at 10%. Estimated regressions control for age, age-squared, female, black, Hispanic, Asian American, core grade point average, predicted SAT test score, father's and mother's years of education, parent earned a bachelor's degree, parents married, parents' income, income-squared, number of siblings, Catholic, other Christian, other religion, non-English speaking household, urban high school and rural high school. Long (2008) also controls for percent of adults in the student's zip code who have a bachelor's degree, percent of adults in the student's zip code who have a graduate degree, unemployment in the student's zip code, income per capita in the student's zip code, and an index of high school quality. Application and admission fixed effects follows Dale & Krueger (2002) method whereas application fixed effects matches students on application portolios while ignoring admission outcomes.

Table 7: Nonlinear Effect of College Selectivity on Graduation

Dependent Variable = 1 if Graduates from Four-Year College, 0 Otherwise

Omitted Variable = Avg. SAT of Enrolled College less than 1000

Twins Only

Variable	OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
Median SAT of Enrolled College 1000 - 1099	0.029**	0.021*	0.019	0.014
	(0.012)	(0.013)	(0.022)	(0.022)
Median SAT of Enrolled College 1100 - 1199	0.133***	0.118***	0.105***	0.096***
	(0.013)	(0.013)	(0.022)	(0.023)
Median SAT of Enrolled College 1200 - 1299	0.169***	0.149***	0.156***	0.145***
	(0.015)	(0.016)	(0.025)	(0.025)
Median SAT of Enrolled College Greater Than 1300	0.204***	0.170***	0.172***	0.154***
	(0.018)	(0.020)	(0.029)	(0.030)
Observations	22,016	22,016	22,016	22,016
R-squared	0.174	0.176	0.057	0.058

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Table 8: Heterogeneous Effect of College Selectivity on Graduation
Coefficient Estimates of Median SAT of College Enrolled (100s)
Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise
Each Coefficient Estimate from Separate Regression
Twins Only

	OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
All Twins (obs = 22,016)				
	0.058***	0.052***	0.052***	0.048***
	(0.004)	(0.004)	(0.006)	(0.007)
Male Only (obs = 7,220)				
	0.057***	0.049***	0.063***	0.061***
	(0.007)	(0.008)	(0.011)	(0.011)
Female Only (obs = 9,340)				
	0.054***	0.049***	0.034***	0.032***
	(0.006)	(0.007)	(0.010)	(0.011)
Mixed Gender (obs = 5,456)				
	0.066***	0.059***	0.064***	0.056***
	(0.008)	(0.008)	(0.012)	(0.012)
White Students (obs = 14,314)				
	0.054***	0.046***	0.049***	0.046***
	(0.005)	(0.006)	(0.008)	(0.009)
Black and Hispanic Students (obs = 2,974)				
	0.045***	0.040***	0.036**	0.030**
	(0.009)	(0.010)	(0.014)	(0.015)
Parents' Income Less Than $50,000 (obs = 4,248)				
	0.040***	0.042***	0.039***	0.038**
	(0.008)	(0.009)	(0.015)	(0.016)
Parents' Income More Than $50,000 (obs = 9,436)				
	0.053***	0.046***	0.049***	0.044***
	(0.006)	(0.007)	(0.011)	(0.012)
Urban High School (obs = 5,792)				
	0.062***	0.056***	0.039***	0.039***
	(0.007)	(0.008)	(0.012)	(0.012)
Small City/Town High School (obs = 3,220)				
	0.033***	0.035***	0.039**	0.036**
	(0.011)	(0.012)	(0.016)	(0.017)
Suburban High School (obs = 9,256)				
	0.064***	0.057***	0.069***	0.066***
	(0.006)	(0.007)	(0.010)	(0.010)
Rural High School (obs = 2,620)				
	0.055***	0.044***	0.049**	0.041**
	(0.012)	(0.013)	(0.019)	(0.021)

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Table 9: Effect of College Selectivity on Longer Graduation Rates
Coefficient Estimates of Median SAT of College Enrolled (100s)
Each Coefficient Estimate from Separate Regression
Twins Only

OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise (obs = 22,016)			
0.058***	0.052***	0.052***	0.048***
(0.004)	(0.004)	(0.006)	(0.007)
Dependent Variable = 1 if Graduates in Five Years, 0 Otherwise (obs = 14,709)			
0.061***	0.062***	0.059***	0.058***
(0.005)	(0.005)	(0.008)	(0.008)
Dependent Variable = 1 if Graduates in Six Years, 0 Otherwise (obs =7,134)			
0.059***	0.065***	0.076***	0.078***
(0.007)	(0.008)	(0.011)	(0.012)

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Table 10: Effect of Undermatch/Overmatch on Graduation

Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise

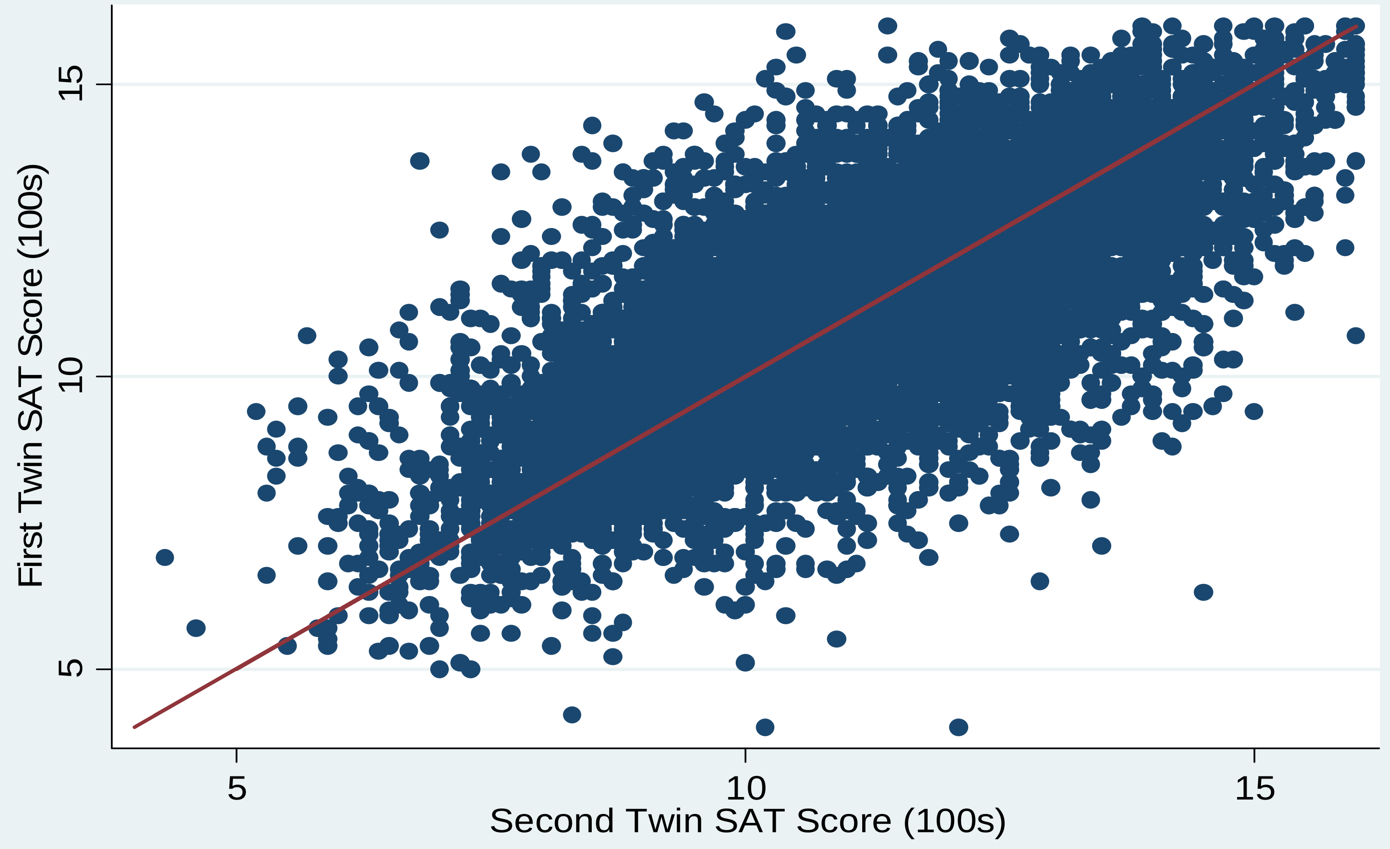
Each Coefficient Estimate from Separate Regression

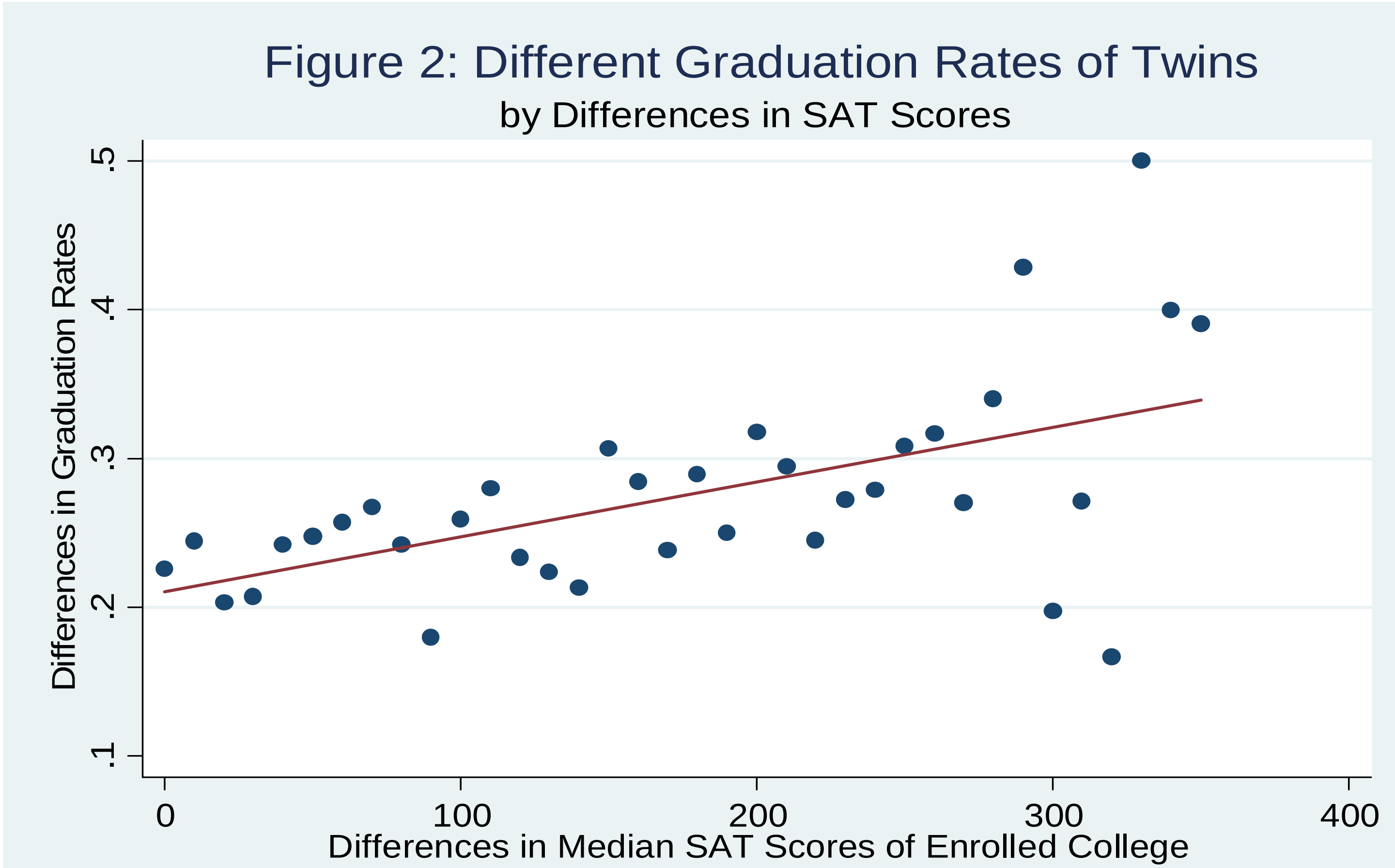
Twins Only

Binary Indicator of Undermatch/Overmatch:	OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
SAT Score 50 Points Greater Than Median at College Enrolled	-0.095***	-0.076***	-0.062***	-0.056***
	(0.009)	(0.009)	(0.012)	(0.012)
SAT Score 100 Points Greater Than Median at College Enrolled	-0.114***	-0.095***	-0.078***	-0.072***
	(0.010)	(0.010)	(0.013)	(0.013)
SAT Score 200 Points Greater Than Median at College Enrolled	-0.095***	-0.070***	-0.083***	-0.076***
	(0.016)	(0.016)	(0.020)	(0.020)
SAT Score 50 Points Less Than Median at College Enrolled	0.084***	0.066***	0.044***	0.037***
	(0.009)	(0.009)	(0.011)	(0.011)
SAT Score 100 Points Less Than Median at College Enrolled	0.074***	0.055***	0.020*	0.012
	(0.009)	(0.009)	(0.012)	(0.012)
SAT Score 200 Points Less Than Median at College Enrolled	0.058***	0.040***	0.034**	0.026*
	(0.012)	(0.012)	(0.016)	(0.016)

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Figure 1: SAT Scores of Twins





Graduation rates are the average differences between the twins in a 10 SAT score bin.

Appendix 1: Nonlinear Effect of College Selectivity on Graduation

Dependent Variable = 1 if Graduates from Four-Year College, 0 Otherwise

Omitted Variable = Avg. SAT of Enrolled College less than 1000

2004 Cohort of Twins Only

Variable	Graduates in Four Years	Graduates in Five Years	Graduates in Six Years
Median SAT of Enrolled College 1000 - 1099	-0.041	0.030	0.030
	(0.038)	(0.043)	(0.043)
Median SAT of Enrolled College 1100 - 1199	0.130***	0.211***	0.211***
	(0.040)	(0.043)	(0.044)
Median SAT of Enrolled College 1200 - 1299	0.169***	0.250***	0.250***
	(0.046)	(0.049)	(0.049)
Median SAT of Enrolled College Greater Than 1300	0.146***	0.239***	0.239***
	(0.052)	(0.056)	(0.049)
Observations	7,134	7,134	7,134
R-squared	0.072	0.061	0.056

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.

Appendix 2: Effect of College Selectivity on Graduation by Cohort
Coefficient Estimates of Median SAT of College Enrolled (100s)
Dependent Variable = 1 if Graduates in Four Years, 0 Otherwise
Each Coefficient Estimate from Separate Regression
Twins Only

OLS	OLS with Portfolio Controls	Twin FE	Twin FE with Portfolio Controls
All Twins (obs = 22,016)			
0.058***	0.052***	0.052***	0.048***
(0.004)	(0.004)	(0.006)	(0.007)
2004 Cohort (obs = 7,134)			
0.073***	0.068***	0.066***	0.064***
(0.007)	(0.008)	(0.011)	(0.012)
2006 Cohort (obs = 7,574)			
0.056***	0.051***	0.055***	0.050***
(0.007)	(0.008)	(0.010)	(0.011)
2007 Cohort (obs = 7,308)			
0.047***	0.038***	0.038***	0.032***
(0.007)	(0.008)	(0.011)	(0.011)

Notes: Standard errors in parentheses and clustered at the twin level. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for SAT score, number of SAT2's taken, dummies for number of AP tests taken and GPA, and a male dummy. OLS regressions control for income, ethnicity, state residence, native language, citizenship, father's and mother's education, county unemployment rate and education, high school percent free and reduced price lunch, and year. Application portfolio controls include number of Score Sends and the minimum, mean, and maximum median SAT score of enrolled students for the colleges in the portfolio of Score Sends.