Do Good High Schools Produce Good College Students?
Early Evidence from New York City

Hella Bel Hadj Amor*
Research Scientist, Institute for Education and Social Policy,
New York University

Amy Ellen Schwartz
Director, Institute for Education and Social Policy
Professor of Public Policy, Education, and Economics, Steinhardt School of Education and Wagner Graduate School of Public Service
New York University

Leanna Stiefel
Professor of Economics, Wagner Graduate School of Public Service
New York University

Abstract: We examine variation in high school and college outcomes across New York City public high schools. Using data on 80,000 students who entered high school in 1998 and following them into the City University of New York, we investigate whether schools that produce successful high school students also produce successful college students. We also explore differences in performance across sex, race, and immigration, and we briefly explore selection issues. Specifically, we estimate student-level regressions with school fixed effects, controlling for student characteristics, to identify better and worse performing schools based on state mandated exams, graduation and college performance.

Classification Codes: I20; I21; I23
Key Words: High school, College, Fixed effects, Value added

Do Good High Schools Produce Good College Students?
Early Evidence from New York City

I. Introduction

While education has long been a pathway to economic success, since the 1980s the premium in earnings for college versus high school education has grown almost continuously. Yet in order to enter -- let alone *succeed* -- in college, a student must earn a high school diploma and gain adequate high school preparation. This pivotal role for high schools in determining enrollment and success in college is receiving increasing attention from researchers as well as policymakers. While much of the attention focuses on using testing for high school accountability, there is increasing interest in the relationship between high schools, college going and success.

As an example, although the *No Child Left Behind Act* of 2001 (NCLB) currently requires testing only in grades 3-8 and one high school grade, in January 2005 President Bush announced his desire to expand the law to include more high school tests as well as high school graduation rates (e.g., Kornblut, 2005). Perhaps more interesting, the state of Michigan recently decided to replace the state high school exams with the American College Test (ACT) plus five other subject exams that "would be accepted by colleges and universities for entrance and placement purposes."

In New York State, students will soon be required to pass exams in five subject areas (the long-standing "Regents'' exams") in order to graduate from high school. In the past, Regents' exams were taken by far fewer than most of the states' high school students, and those taking the exams were probably the most academically accomplished as well as the most likely to attend college. Since in the near future all students will be required to pass Regents' exams to graduate from high school, the Board of Regents' seems to intend to broaden the pool of college-ready students.

Even if more policies focus on the success of schools at turning 8th graders into graduates or into graduates ready for college, insufficient attention is being paid to whether these "successes" actually do translate into successful college students. While some researchers have looked at the high school characteristics of individual students (e.g. grade point average) when assessing college success, very few have studied how high schools perform as organizations *per se*. Yet in a world of increasingly stringent accountability for organizational outcomes, such research is crucial in order to understand if there is conflict or congruence between the various missions high schools are asked to achieve, and more specifically between the ways these missions are measured. The primary purpose of this paper is to begin to fill this gap by examining the extent of variation in high school and college outcomes across New York City public high schools and to investigate whether high schools that produce students who do well on high school measures of success also produce students who do well on college measures of success.

Note that this is early evidence because the requirement to obtain five passing grades on Regents' examinations is not yet fully implemented and because we do not yet have college persistence or graduation information for our sample of students. Nevertheless, it is important to assess whether success is likely from early evidence and to ascertain whether there are conflicts between traditional measures of successful high schools (such as graduation rates or passing scores on high school exams) and less traditional measures of successful high schools (such as college application or matriculation or grades).

We use data on 80,000 New York City (NYC) public high school students who entered ninth grade in 1998, follow them through their high school years and, where relevant, into the City University of New York (CUNY). We investigate whether schools with high graduation rates and test scores also produce students who earn high grade point averages (GPAs) at CUNY. Specifically, we estimate student-level regressions with school fixed effects, controlling for student characteristics, to rank high schools along a number of different indicators, including state mandated high school exams, graduation rates, Scholastic Aptitude Tests (SAT) scores, and college performance. The school fixed effects capture the average value added by the school to its student performance. We then compare high school success along these different dimensions. Results will be useful in devising accountability measures for high schools. If high schools are ranked similarly on all indicators of success, then the need to evaluate tradeoffs may be mitigated. Otherwise, policymakers will need to better align measures of success or at least be attentive to the choices of outputs for which schools are held accountable.

The paper is organized as follows. In section II, we review the relevant literature to provide background for the study. In sections III and IV, we describe the data and the methodology, respectively. In section V, we present results on the relationships between various indicators of high school performance and in section VI we conclude with a summary and discussion of the implications of our results for policies on accountability.

II. Literature Review

The transition from high school to college

There is not a great deal of research on the relationship between how well high schools perform and how successful their students are in college. Nonetheless, within this scant literature, two strands emerge, one that examines the relationships between high school outcomes and college outcomes, and another that examines the relationship between specific policies and college entrance exams and outcomes.

In the first strand are papers that estimate straightforward multivariate models of college GPAs, controlling for high school outcomes (e.g., high school GPA, class rank, receipt of general equivalency diploma (GED), and SAT scores), and a limited set of student characteristics. For example, Cohn et al. (2004) estimate such models with a sample of about 500 South Carolina students, focusing on race and sex differences in the likelihood of success in college, eligibility for statewide scholarships, and retention of scholarships after enrollment. Bailey and Weininger (2002) compare credits earned and college attainment of foreign-born and native-born students, while Horowitz and Spector (2005) examine the impact of private versus public high school attendance on the college GPAs of 15,000 Indiana undergraduates.

A second strand of literature analyzes the effect of policies, such as state-required graduation exams (e.g., Marchant and Paulson, 2005) or state high-stakes tests (e.g., Ehlert and Podgursky, 2005), on college entrance exams and outcomes. Again using straightforward multivariate regression, Marchant and Paulson (2005) model graduation rates and SAT scores, at the state and student levels, controlling for a few demographic characteristics and, in the SAT models, for high school GPA and an indicator for whether the state requires a standardized test for graduation. A different approach is taken in Ehlert and Podgursky (2005), who assess whether low-stakes state high school performance assessments are reliable measures of true school performance. Their methodology differs entirely from the other papers, however, as the authors eschew regression analysis and instead use an array of correlations between performance levels on proficiency exams and ACT scores, as well as distributions of these measures and of college enrollment, credits and GPA.

Three important differences exist between our study and those in the literature. First, our interest lies in high schools *per se* rather than individual students, and more specifically, in whether high schools with high value added on high school measures of success also have high value added on college measures of success. Thus, we not only estimate models of college GPAs (and SATs), but also models of high school outcomes (graduation and performance on New York State high school Regents' exams) and we do this with a focus on high school performance rather than individual student performance. Second, in some models, we do not control for students' high school outcomes, but rather for the quality of students at high school intake, using their 8th grade test scores, thus allowing us to compare the "total" high school effect with one that is mediated by intermediate student high school outcomes. Third, for readers who wish to analyze the effect of individual student characteristics, we drastically reduce omitted variable bias on coefficients of those characteristics by including high school fixed effects and by controlling for a much wider array of student demographics than most papers cited, except perhaps Bailey and Weininger (1999). We include sex (as in Bailey and Weininger, 1999, Cohn et. al., 2004 and Horowitz and Spector, 2005), race (as in Bailey and Weininger, 1999, Cohn et. al., 2004 and Marchant and Paulson, 2005), age (as in Bailey and Weininger, 1999, Horowitz and Spector, 2005), poverty status (Bailey and Weininger, 1999, use measures of household income), immigrant status (the variable of interest in Bailey and Weininger, 1999), as well as language ability and special education status.[1]

Our fixed effects specification follows Betts and Morell (1999), who include school fixed (or random) effects in a model of college GPA, although their focus is not on high school effects but rather on the other variables in the models such as family background, high school resources and peer characteristics. Interestingly, the authors obtain similar results when they substitute random effects for school fixed effects. In addition, while the authors indicate the possible presence of selection bias (because some schools send their best students to the University of California at San Diego, while other schools send their second-best students), they argue that their study contributes to the literature because it is an early study that includes student, school

[1] Poverty and special education are included in Marchant and Paulson (2005) in the state-level models.

and neighborhood characteristics.[2] In the same vein, our paper serves as an early study of the congruence between high school success on high school and on college outcomes.

Successful high schools

By estimating models of both high school and college success, we draw on and expand the literature on high school effectiveness, as well as literature linking high schools to college. Using our unique New York City student-level database, we add to the high school literature, which has drawn in large part on the National Education Longitudinal Study of 1998 (NELS:88). For example, Goldhaber and Brewer (1997) estimate an education production function, modeling the math achievement of 5,000 tenth graders as a function of individual and family characteristics, as well as school, teacher and classroom variables. They expand their analyses by also estimating models with random and fixed effects (both school and teacher), and regressing teacher effects on teacher characteristics.

Other studies start with more complicated models, nesting students within schools in a hierarchical linear framework. Lee and Smith (1997), for example, are ultimately interested in the effect of high school size on the average test score gains of students in schools with varying representations of poor and minority students. Warren and Edwards (2003) use a similar setup to investigate the effects of high-stakes graduation test requirements on high school attainment.

In this work, we estimate high school fixed effect models of student performance and use the fixed effect coefficients to compare high school performance across indicators of high school and college success. In other words, we compare the average value added of high schools to assess the extent to which high schools that exhibit high value added on the high school outcomes are similarly successful on the college outcomes.

Performance measurement

There are two common ways that economists measure organization performance in general and school performance in particular. In previous work, we have labeled one method "adjusted performance measures" (APMs). To measure the performance of elementary and middle schools in New York City and Ohio, for example, Stiefel, Schwartz, Bel Hadj Amor and Kim (2005) use residuals from school-level, multiple-regression equations in which raw performance indicators are the dependent variables. We call these APMs because they adjust the

[2] See Rothstein (2004) for an in-depth discussion of selection bias in modeling the relationships between college GPA and SATs.

output measure (e.g. the percentage of students passing tests) for student and school characteristics beyond the control of an individual school, such as student educational need and school resources. A lagged performance measure is included as an independent variable to approximate the value added of achievement over the school year.

While using residuals to measure and compare an organization's performance, especially for schools, is not new (other examples include Gramlich, 1976; Ladd and Walsh, 2002; Rubenstein, Schwartz and Stiefel, 1998; Stiefel, Rubenstein and Schwartz, 1999, 2003), few papers perform such analyses for high schools. Rumberger and Palardy (in press) is an exception. These authors investigate the relationship between several different indicators of high school performance, in an effort to answer a question similar to ours: Are schools that are effective in raising test scores also effective in reducing dropout rates, transfer rates and attrition rates? The authors use the residuals from a set of student-level models that control for students characteristics and a rich set of school characteristics to measure and compare school effectiveness.

School fixed effects are an alternative way to measure and compare organizational effectiveness (e.g., Bartelsman and Doms, 2000; Stiefel, Bel Hadj Amor and Schwartz, 2005). They can be included in student-level models of school performance as long as the included school-level variables are time-variant. Fixed effects are superior measures because residuals include an array of random components, while fixed effects capture the schools' time-invariant, unobserved characteristics, and can be interpreted as the schools' average value added to their students' performance.[3] In this paper, we use the school fixed effects from student-level regressions to rank high schools.

III. Data

New York City Students

We use data on a high school cohort of approximately 80,000 students who were expected to graduate from the New York City high schools in 2001. The New York City Department of Education (NYCDOE) provided data on high school students, their demographic

[3] O course, time invariant unobserved characteristics may represent more than the school's input *per se*. For example, if parental selection of high schools is unvarying or consistent across schools, and parental characteristics of individual students are not adequately controlled in the fixed effect regression, then the fixed effects will reflect parental selection as well as the school's contribution to value added. We thank Sarah Turner for this insight.

characteristics (including race, sex and immigrant status), their performance on New York State Regents' high school exams, Regents' Competency Tests, 8th grade reading and math tests, the schools they attended throughout their high school career, and whether they graduated high school within four years, and if so, with what type of degree.[4] CUNY provided additional information on high school graduates, their SAT scores, as well as application and enrollment information, including GPAs.

The analysis sample includes nearly 50,000 of the 80,000 students. We exclude 14,000 students who were at one point part of the class of 2001, but who were discharged from New York City high schools (usually to another district or a private school) before graduation.[5] In addition, 2,700 students graduated with a GED and 260 with a special education certificate. The behavior of these latter two groups is unlikely to be captured by the same models that describe the students who have a regular high school experience, i.e. students who, after four years, are still enrolled, graduate with local or Regents' diplomas, or drop out. In addition, in very small schools, the averages of the variables of interest will be very sensitive to the presence of outliers. To avoid the resulting potentially skewed results, we limit the sample to students in schools with more than ten students.[6] Accordingly, the models restricted to CUNY *enrollees* use students in high schools with more than ten enrollees. Further, estimating reliable fixed effects for high schools requires a minimum number of students in each high school who have data on the dependent variables of interest. We drop from the remaining sample students from high schools with five or fewer students who enroll in CUNY and have a GPA, SAT scores and English and math Regents' scores. The resulting sample has 50,494 students in 148 high schools. This sample includes 31,453 high school graduates and 13,342 CUNY enrollees. Descriptive statistics are presented in Table 1.

TABLE 1 ABOUT HERE

The NYC high school student population in our sample is non-white and poor (Table 1 columns (1), (2) and (3)). Indeed, only 18% of the students are white, while about a third each is

[4] New York State currently awards two high school diplomas, a Regents' and a local. The two differ in that receiving a Regents' diploma requires taking additional credits and passing a larger number of Regents' exams.
[5] Students are discharged when they leave the NYC public high school system. Most students in this situation (71%) are those who leave the city. Others are admitted to a parochial or private school or a high school equivalency program, are home-schooled or institutionalized. Some leave the system when they reach 21. Suspension, expulsion and death are other reasons for discharge.
[6] This eliminates 22 schools enrolling a total of 74 students.

black and Hispanic and 16% are Asian. Seventy-eight percent of the students for whom we have 8th grade school lunch data (69% of the sample) are eligible for a free lunch, and an additional 6% are eligible for a reduced-price lunch. There are fewer males than females (48% and 52% respectively) and, while the average age of the students in 2001 is about 17, age varies from as low as 14 to as high as 24 years old. The diversity in the student population is evident in its racial distribution and in other attributes as well: 23% of the students were born in a foreign country and 50% speak a language other than English at home. A much smaller percentage, 12%, is English language learners. This is not entirely surprising, as the vast majority of the students (91%) entered the NYC school system before the 9th grade.[7] Five percent of the students in the sample received part-time special education services when they were in 8th grade.

There is a wide range in student performance in our sample. While average performance in reading and math on the 8th grade tests is about 0.15 standard deviations above the mean across all students in the original dataset, it deviates from the mean by almost 3 standard deviations in both directions. Disparities are large in high school as well, and scores on the English and math Regents' exams ranges from 1 to 100, averaging about 68. Students usually took the English Regents' exam in 2000 and the math exam in 1998 or later. Only two-thirds of the students graduated high school after fours years (62%).[8]

Enrollees in CUNY look somewhat different from the students as a whole (Table 1 columns (4), (5) and (6)). The share of female is larger (59%), as are the shares of white and Asian students (23% and 20% respectively). The difference in poverty is small, with 83% of poor and near-poor students among the enrollees, versus 84% for the whole sample. There are higher shares, among the enrollees, of students who are foreign-born (27%), were English-language learners in high school (13%) and speak a language other than English at home (55%). This suggests that these students partly overcome language and cultural handicaps, if any, by the time they reach college. Differences in average performance across the two groups are striking however: the average reading and math 8th grade z-scores are 0.32 and 0.35 (vs. 0.15 and 0.16 in the whole sample). While the average English Regents' score is only about one point higher than it is with the whole sample (68.90), the average math Regents' score is four points higher than it

[7] Note that time of entry into the system is unknown for 9% of the students.
[8] This is a higher percentage than that for all New York City students, which is below mid 50 percent, due to the omission of schools and students from our sample as described in previous paragraphs.

is with the whole sample (73.06). Close to the full range of college GPAs is observed, with an average at 2.57.

There is some evidence that the highest-achieving graduates do not enroll in CUNY, as average performance is lower for the enrollees than it is for the graduates on the 8th grade reading test (0.32 vs. 0.46) and the 8th grade math test (0.35 vs. 0.49).

New York City High Schools

To illustrate the great variety in New York City high schools, we aggregate the student-level data to the school level and complement it with data from the NYCDOE Annual School Report (ASR) and School-Based Expenditure Report (SBER) databases. The ASR database includes student demographic and performance information as well as teacher characteristics. The SBER database provides expenditure data for each school.[9] Table 2 provides descriptive statistics on the sample schools.

TABLE 2 ABOUT HERE

Table 2 shows that the 148 high schools in the sample vary widely in their characteristics. While the average school enrolls a little over 1,750 students, there are very small schools (151 students) and very large schools (4,631 students). There is also much variety in student performance, as attested by the wide ranges in the Regents' scores and percentage of students who took the exam. None take them in some schools while all of the students take them in others. While about 57% of the students graduate within four years, some schools graduate all of their students in that timeframe while that number is as low as 14% in other schools. Two-thirds of the students apply to CUNY on average, with much variation in this percentage and in the applicants' SAT scores. In the average school, 42% of the students enroll in CUNY, and this percentage can be as low as 6% and as high as 66%.

There is also great variation in the demographic characteristics of the students, although the average school enrolls a majority of non-white students and poor students. Some schools have populations that are entirely foreign-born, and others have no students who are new to the system. As expected, some high schools receive much higher-performing incoming classes than others; indeed, the 8th grade z-scores range from less than one to over two.

---

[9] Note that the ASRs and SBERs provide data for the whole school, while the aggregated student-level data are limited to the students in the 2001 cohort, i.e., students who were expected to graduate in 2001.

Teacher characteristics vary greatly across schools as well and, accordingly, so does spending. The average school spends a little over $10,000 per pupil in 2001-02, and 71% of its teachers have taught in the school for over two years, 61% have been teaching for more than five years, 84% are licensed and 80% have at least a Masters' degree. Yet some of these percentages are lower than 50% in some schools and reach almost 100% in others.

IV. Framework and Empirical Model

High schools with high value added on the high school outcomes may differ from those with high value added on the college outcomes, for two reasons. First, success on high school exams may not be a good indicator of future success in college. Indeed, indicators of high school and college success may measure different abilities and sets of knowledge and have different aims. Thus, high Regents' exams scores, for example, may not imply high SAT scores. This feature may result in unintended consequences for accountability, if the accountability system is based on one test rather than another.[10]

Second, the set of high schools that excel at getting students into college may be a different set of schools than those that do not (e.g. schools that prepare students for the job market). Thus it is critical to define what a good high school is supposed to do. Since not all students proceed to college, a good school may be defined as one whose students obtain high scores on high school level standardized tests. Such a restrictive definition, however, clearly disadvantages schools whose entering students (through no fault of the high school) are low performers. In this case, a more appropriate definition may take into account that a good high school not only produces students with high test scores, but also enables students to achieve large gains. In such a school, low-performing students who come in from the 8th grade can eventually achieve higher than expected state standardized test scores at the high school level, and graduate with a reasonably high probability. This definition is also more likely to reflect the concerns of policymakers and other stakeholders.

Using this "value-added" definition of success reinforces the possibility that schools that produce high school success and those that produce college success differ. Indeed, what would happen if high schools with comparable average high school test scores, regardless of the level of

---

[10] As noted in the introduction, the Governor of Michigan has announced that all high school students will take the ACT test as a prerequisite for graduation, in part to avoid this disjuncture in measurement.

their entering classes, were compared based on how well their students perform in college? High schools that achieve a large value-added in high school may deliver students for whom achieving a large value-added in college is difficult, while other schools may send students whose high school value-added is more marginal, for whom the potential for gains is yet to be depleted. All of these schools are successful on the high school outcomes, but only the second set would be successful on the college outcomes.

The underlying conceptual framework for our analyses is grounded in adjusted performance measures. In order to discern the true contribution of high schools to their students' success, we first purge confounding factors such as the characteristics of the students and the preparation the students bring with them from home and from middle school. The centerpiece of our empirical work is a set of student-level models that remove the influence of student demographic characteristics and prior performance from student outcomes and produce a set of school indicators that capture the remaining contribution of schools.

First, we model the performance of CUNY enrollees at CUNY (their first semester GPA), in a model of adjusted performance:

(1) $$Performance_{ij} = \alpha_0 + \alpha_1 Student_{ij} + \alpha_2 Score8_{ij} + \varepsilon_{ij}$$

where i indexes students and j high schools. $Performance_{ij}$ represents first semester CUNY GPA.[11] $Student_{ij}$ represents a set of student characteristics, including sex, race and immigrant status. Including the student's performance on 8th grade exams creates a value-added specification of the model, providing some control for differences in student ability and $Score8_{ij}$ represents performance on these tests.[12] α_0 is the intercept, α_1 and α_2 are slopes that capture the impact of the corresponding variables on the outcome, and ε_{ij} is an error term with the usual properties.

Schools are introduced in this model by adding a set of school fixed effects, η_j, and this is the model we estimate:

(2) $$Performance_{ij} = \alpha_0 + \alpha_1 Student_{ij} + \alpha_2 Score8_{ij} + \eta_j + \varepsilon_{ij}$$

[11] We define as the first semester, the first semester when a student is enrolled in CUNY, takes credits and earns a GPA.

[12] These are only available for students who were in the NYC public school system in the 8th grade (almost 70% of the students). Missing flags are used to maintain the sample size.

The school fixed effects capture unobserved characteristics of the schools that affect value added in student outcomes. More specifically, each fixed effect represents the contribution of each school to the student outcome, relative to a reference school, beyond the composition of the student body. The larger the fixed effect, the greater the contribution of that particular school. Thus, school contribution can be measured and compared across schools based on the size of the fixed effects.

Student characteristics include student age and indicators for whether the student is female; black, Hispanic or Asian (white is the left-out category); an English language learner; whether they were born in the United States, whether they entered the NYC public school system before the 9^{th} grade and whether they speak English at home. We use a substitute indicator, when available, for two important student characteristics not reported in high school but reported for 8^{th} graders: student eligibility for free or reduced-price lunch and whether students receive part-time special education services. Z-scores on the 8^{th} grade reading and math standardized tests (CTB/McGraw Hill Test of Basic Skills for English language assessment; California Achievement Test (CAT) for math assessment), as well as the 8^{th} grade score on the Language Assessment Battery (LAB), capture past performance. Indicators for missing data are included for race, time of entry into the system, and all 8^{th} grade variables, so as to maintain sample size. Finally, the model includes the number of semesters between high school graduation and CUNY enrollment.

Model (2) is estimated for both college and high school outcomes. College outcome is measured by GPA and high school outcomes include student performance on the state mandated Regents' exams in English and Sequential I, their verbal and math SAT scores, and an indicator of whether a student graduates.[13]

Results from equation (2) for each outcome provide us with sets of school fixed effects, one for each college and high school outcome. We examine the strength of the relationship between the sets, to determine whether schools that contribute the most to one outcome also contribute the most to the other outcomes.

[13] These models do not include the number of semesters between high school graduation and CUNY enrollment.

V. Empirical results

Raw measures are poor indicators of performance

Adjusting performance for student characteristics and prior performance is a necessary step in evaluating school success. In fact, holding schools accountable for success based on raw measures of performance, such as average graduation rates or Regents' scores, and holding them accountable based on adjusted measures (or value added), i.e. on fixed effects from models of graduation rates or Regents' scores that adjust for student characteristics and prior performance, would lead to very different sets of schools being deemed successful.

To calculate adjusted, fixed effect measures, we estimate models of high school performance, controlling for student characteristics and prior performance. Table 3shows results for models of the likelihood of graduating and of the English Regents' scores. Coefficients are similar in sign across models, and specific results are described in the Appendix.[14]

TABLE 3 ABOUT HERE

Using our estimated value-added fixed effects in more detail, we find that even though correlations between the raw and value-added measures of the same outcome are high (close to 0.90), correlations across measures of different high school outcomes are much lower when the value-added measures are compared than when raw measures are compared. This suggests that although the same high schools may at first appear do be doing well (poorly) on all high school measures of success, once measures are adjusted, there are differences that depend on the choice of the indicator of success.

Specifically, comparing raw outcomes suggests that schools that are successful on one high school outcome are generally successful on the others (Table 4, top panel). For example, schools with high Regents' exam scores also have high graduation rates: the Pearson correlations between the outcomes are 0.75 and 0.78, respectively, for English and Sequential 1 (math) Regents' and graduation. This pattern also describes the relationship between the English and Sequential I Regents' scores (r = 0.79). Overall, these results suggest that holding schools accountable for one raw high school outcome would lead to roughly the same set of successful schools as holding schools accountable for another raw high school outcome.

---

[14] SAT scores are available for CUNY applicants only, so we cannot estimate these models for all students.

TABLE 4 ABOUT HERE

Caution is warranted, however. Value-added measures of high school outcomes are much less correlated than unadjusted measures (Table 4, bottom panel). Indeed, the correlations among the value-added fixed effects from the graduation and Regents' models fall to around 0.50 (versus close to 0.80 for the unadjusted, raw measures). Thus it is important to consider a range of high school outcomes when assessing the success of a school, and it is critical that these outcomes be adjusted for factors outside the control of schools.

High school and college value-added outcomes of enrollees are not highly correlated

Next, we use value-added measures to evaluate how successful high schools are with the students they send to CUNY, on both their high school and college outcomes. Once again, the value-added measures are estimated in a series of student-level regressions with school fixed effects, controlling for student characteristics and prior performance and estimated using a sample restricted to CUNY enrollees (Table 5). Results of these regressions are described in detail in the Appendix.

TABLE 5 ABOUT HERE

Most of the correlations between the high school and college value-added outcomes of the enrollees are low (Table 6). The GPA fixed effects are somewhat highly correlated with the SAT fixed effects (0.47 with verbal and 0.41 with math), but the correlations drop sharply when the GPA fixed effects are compared to the Regents' scores fixed effects (0.25 with the English Regents' and zero with the Sequential I Regents').

TABLE 6 ABOUT HERE

Cross-tabulations (not reported) of the value-added outcomes divided into three groups of roughly the same size confirm these results. In cross-tabulations of the high school measures of outcome, whether we are looking at high school students or CUNY enrollees, or across the two samples, close to 50% of the schools, and sometimes as many as 60% of the schools are on the diagonal, i.e. they are in the bottom for both measures, or in the middle for both measures, or at the top for both measures (the numbers are even higher if we examine cross-tabulations of the same outcome across high school students and CUNY enrollees; they can be higher than 90% of the schools). In general, less than 15% of the schools are at the opposite extremes, i.e. are at the top for one measure and the bottom for the other. On the other hand, in cross-tabulations that compare high school to college outcomes, that is fixed effects for the high school outcomes of the enrollees (or the high school students) and for the GPA of enrollees, the percentage of the schools on the diagonal tends to be well below 45% (most often, 36-38%), while the percentage of the schools at the extremes tends to be at least 15% and sometimes over 25%. The SAT and college GPA fixed effects exhibit the highest percentages on the diagonals.

Focusing on the enrollees does not bias the results

High schools may experience different levels of success with the high school outcomes than they do with the college outcomes if the students who enroll in college are systematically different from those who do not. We find, however, that models of high school Regents' performance estimated with all high school students are not very different from those estimated with the enrollees (the Regents' are the two outcomes that are available for both subgroups).[15]

Further, Table 7 indicates that the correlation between the fixed effects from the Regents' performance models is very high, 0.89 (using the Sequential I Regents' leads to the same conclusions). This suggests that limiting the sample to the enrollees will not yield misleading conclusions.[16]

TABLE 7 ABOUT HERE

Changing college performance models to control for high school success leaves high school rankings materially unaffected

So far, we have compared the fixed effects from models of high school and college performance that control for the performance of students at intake into high school, using their

[15] Comparing the results from the English Regents' regressions for the enrollees and the high school students, we find only a few cases where a variable that does not affect the enrollees affects high school students. For example, for the latter, the Regents' scores decrease as age increases; they are lower for Native Americans and, surprisingly, for students who speak English at home; and students who entered the system after 9th grade do better that the rest of the students.

[16] Note also from the last column in this table that high value-added on the high school outcomes when all students are considered seems unrelated to value-added on the GPA of enrollees. This may reflect the fact that students who choose to enroll in CUNY are not a random sample of the high school population, where different students enroll in more or less selective colleges, or choose not to enroll.

8th grade test scores. Next, we examine college performance controlling for the high school performance of students (Table 8). Regression results are very similar to those obtained when controlling for 8th grade performance, the one exception being that ELL students have a higher GPA than non-ELL students (this was only true for native-born students when we were controlling for 8th grade performance). The higher the high school scores, the higher the GPA. The fixed effects from this model are extremely highly correlated to the fixed effects from the model of GPA that controls only for 8th grade performance (0.97)

TABLE 8 ABOUT HERE

VI. Summary and Discussion of Results

Summary

The roles of U.S. high schools are changing. Toward the middle of the 20th century, American high schools were meant to provide an opportunity to achieve a high school education, but even students who did not or could not avail themselves of this opportunity (that is dropouts) were able to make a decent living. More recently, the public and policymakers seem to have higher ambitions for public high schools, namely to graduate the vast majority of youth. In addition, currently there is considerable discussion of the role of high schools in producing college ready students, especially since there is some evidence that youths as old as 8th graders overwhelmingly aim to complete college, and additional labor market evidence that even a high school diploma will not land a good job.

Higher ambitions combined with increasingly more stringent accountability standards at the high school level make it important to ascertain if the common measures of high school success are consistent with measures of college success. This paper has presented some early evidence on this issue. We estimate regression-adjusted value-added student outcomes, with high school fixed effects, for a variety of outcomes at both the high school and college levels. In the paper, we report on a subset of these outcomes and find four results.

First, raw unadjusted outcomes of high school success at the high school level (Regents' test results, SAT scores, and graduation rates) correlate well. Second, adjusted, value-added outcomes of high school success at the high school level (Regents' test results, SAT scores, and graduation rates) do not correlate well. Since value-added, adjusted measures are clearly superior to raw outcomes as ways to evaluate the role of a school, this is an unsettling result. Third, when college outcomes (that is college GPAs) are compared to high school outcomes, the correlations with high school SAT scores are around .5, but all other correlations are considerably lower. Fourth, these results do not seem to be a product of selection issues, whereby some schools produce college students and others do not.

Discussion

If policymakers want to measure high school success at producing college students who succeed, they should consider using the already developed and available SAT or ACT test, as has the state of Michigan. A state could pay for all students to take the test and even establish its own criteria for graduation if it desired. This route would use fewer state resources than developing its own tests and would be useful to colleges. High school SAT fixed effects are correlated with college success in our early work and, as reported here, are the best of the alternative ways to assess high school contribution to college success.

If, on the other hand, policymakers want even more from high schools, for example that students not only have aptitude for college work, but also that they have knowledge of a body of material such as how to solve for an unknown in an algebraic equation, or the structure of DNA, or significant events in American history, or identification of important writers etc., then additional tests would be needed. But more tests will be more expensive to administer -- to design, implement and grade -- and also more expensive in terms of student, teacher and other effort. Are they worth it?

If we really do not think all students are capable of completing additional years of education beyond high school, then are measures of college success the right ones on which to base our measures of high school success? Are skills for the marketplace and/or technical training perhaps different than those needed for college success? If they are different, should all high school students be expected to accomplish both types of skills? This will certainly be an expensive proposition if many college students are mechanically or electrically skills "challenged" and, conversely, if many mechanics, repair people, and construction workers are college aptitude "challenged." In order to avoid sorting students by type of work preferred or by aptitude for different kinds of work, how many resources are we willing to expend? This is an old question and harks back to questions of when and if "tracking" should occur and whether technical schools can really prepare students for the workplace or if on-the-training is more appropriate. In the U.S., we seem to be headed in the direction of college-ready high school students, but in Europe, this is less true. The advent of stringent accountability for school success may force policymakers and the public to decide if one high school fits all or if there can or should be different kinds of high school education available to students.

The work reported will benefit in the future from additional data on the success of college students, including the majors students choose, their persistence and their graduation rates. In addition, more complete information on college choices that go beyond the CUNY system will be helpful, although 68% of our sample does apply to CUNY and 43% enroll, making the information about CUNY particularly relevant. Knowing the course of study of high school students could be an important measure of high school success since next to the SAT or ACT scores, college admissions offices often claim that the rigor of the course of study predicts admission (and success). Despite all the added information that additional data will provide, however, early evidence indicates that success in college and success on traditional high school indicators are quite different.

Acknowledgments

We thank Sarah Turner, Norm Fruchter, Dorothy Siegel and the research workshop of the Institute for Education and Social Policy, New York University, for helpful suggestions. We are solely responsible for the paper's content.

Appendix: Regression Results

Models of high school performance (likelihood of graduating and English Regents' scores), controlling for student characteristics and prior performance, as reported in Table 3, are similar overall. The relationship with age varies: while the English Regents' score decreases consistently as age increases, the likelihood of graduation decreases as students age to 20, then increases.[17] Asians are more likely to graduate than other students, but their English Regents' scores are lower than those of the other students. The English Regents' are not affected by poverty, and poor students are less likely to graduate than the rest of the students. Immigrant status, the time of entry into the system, and ELL status interact differently in the way they affect the outcomes. One common finding is that ELL students tend to be worse off, especially if they entered the system before 9th grade.

The student-level models with school fixed effects, controlling for student characteristics and prior performance, and estimated using a sample restricted to CUNY enrollees (Table 5) show that females score higher than males on the English Regents', yet they have lower verbal SAT scores than males have, in spite of which they achieve higher GPAs than males do. Age only affects the SAT score, which decreases as students age between 14 and 22, then increases for students between the ages of 22 and 24. Not surprisingly, white students do better than all the other students, and black students do worse than white, Hispanic and Asian students. Asians outperform Hispanics on the GPA and Regents', but Hispanics do better on the SAT. Native Americans have the lowest GPAs, but their Regents' and SATs are no different from those of whites. Interestingly, for the most part, the performance of poor students is on a level with that of non-poor students, except that students who received a reduced-price lunch in the 8th grade have lower GPAs than the rest of the students. While this may suggest that the poor have overcome their academic disadvantage by high schools, it may instead reflect the fact that the free lunch data do not fully capture poverty, and caution should be used in interpreting these coefficients.

Native-born students have lower GPAs than foreign-born students, and those among them who are not English language learners do worse than the rest (-0.134 versus -0.02). Time of entry into the system does not affect a student's GPAs. This may be due to enough time passing before

[17] This is also true of the Sequential I Regents' score model (not reported). Other notable results from this model include the fact that Asians outperform the other students, and so do the poor, surprisingly.

college entrance, since time of entry into the system does affect high school outcomes -- Regents' scores and SATs.

There is an even more complex relationship between the Regents' and verbal SAT scores, and whether students are native-born, ELL and their time of entry into the system. The reference group is foreign-born students who enter the system after the 9th grade and are not ELL. One group only outperforms these students, and they are the native-born students who also entered the system after the 9th grade and are not ELL. One other group has a comparable performance to the reference group, and these are foreign-born students who are not ELL but entered the system before the 9th grade. All the other subgroups of students do worse than the reference group on these measures, with foreign-born ELL students doing particularly poorly. Interestingly, students who speak English at home do better on the SATs but worse on their college GPA. Their Regents' scores are not affected.

Students who received part-time special education services in the 8th grade do worse on the Regents' and SATs, but their GPA is unaffected, suggesting that this handicap has been overcome by the time these students reach college, perhaps thanks to the additional services they received in middle school.

The relationships between the outcomes and most 8th grade test scores are non-linear, suggesting that outcomes increase then decrease with 8th grade scores depending on the levels of those scores. The one simpler relationship suggests that GPA increases consistently as 8th grade math scores increase. The GPA is unaffected by 8th grade LAB scores, while Regents' and SAT scores decrease as LAB scores increase.[18]

[18] Models with other dependent variables (math SAT and Sequential I Regents', available from the authors) yielded only slightly different results.

References

Bailey, T., Weininger, E.B., 2002. Performance, Graduation, and Transfer of Immigrants and Natives in City University of New York Community Colleges. Educational Evaluation and Policy Analysis, 24, 359--377.

Bartelsman, E.J., Doms, M., 2000. Understanding Productivity: Lessons From Longitudinal Microdata. Journal of Economic Literature, 37, 569--594.

Betts, J.R., Morell, D., 1999. The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources, and Peer Group Effects. Journal of Human Resources, 34, 268--293.

Cohn, E., Cohn, S., Balch, D.C., Bradley, J., 2004. Determinants of Undergraduate GPAs: SAT Scores, High-School GPA and High-School Rank. Economics of Education Review, 23, 577--586.

Ehlert, M., Podgursky, M., 2005. What Happens If They Aren't "Proficient?" The Predictive Validity of a High School State Assessment. Paper presented at the Third Research Seminar in Analytic Issues in the Assessment of Student Achievement, The Urban Institute, Washington, D.C. May 2, 2005.

Goldhaber, D.D., Brewer, D.J., 1997. Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity. Journal of Human Resources, 32, 505--523.

Gramlich, E.M., 1976. The New York City Fiscal Crisis: What Happened and What Is to Be Done? American Economic Review, 66, 415--429.

Horowitz, J.B., Spector, L., 2005. Is There a Difference Between Private and Public Education on College Performance? Economics of Education Review, 24, 189--195.

Kornblut, A.E., 2005, January 13. Bush Urges Rigorous High School Testing. The New York Times, 26.

Ladd, H.F., Walsh, R.P., 2002. Implementing Value-Added Measures of School Performance: Getting the Incentives Right. Economics of Education Review, 21, 1--18.

Lee, V.E., Smith, J.B., 1997. High School Size: Which Works Best and for Whom? Educational Evaluation and Policy Analysis, 19, 205--227.

Marchant , G.J., Paulson, S.E., 2005. The Relationship of High School Graduation Exams to Graduation Rates and SAT Scores. Education Policy Analysis Archives, 13. Retrieved August 24, 2005 from http://epaa.asu.edu/epaa/v13n6/.

“New Exam to Put All Michigan Students on Path to College.” Retrieved October 23, 2005 from http://www.michigan.gov.

Rothstein, J.M., 2004. College Performance Predictions and the SAT. Journal of Econometrics, 121, 297--317.

Rubenstein, R., Schwartz, A.E., Stiefel, L., 1998. Conceptual and Empirical Issues in the Measurement of School Efficiency. Proceedings of the 91st Annual Conference on Taxation. Washington, D.C.: National Tax Association, 267--274.

Rumberger, R.W., Palardy, G.J., in press. Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance. American Journal of Education.

Stiefel, L., Bel Hadj Amor, H., Schwartz, A.E., 2005. Best Schools, Worst Schools, and School Efficiency: A Reconciliation and Assessment of Alternative Classification Systems, in: W. J. Fowler (Ed.), Developments in School Finance, 2004. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 81-101.

Stiefel, L., Rubenstein, R., Schwartz, A.E., 2003. Better Than Raw: A Guide to Measuring Organizational Performance With Adjusted Performance Measures. Public Administration Review, 63, 607--615.

Stiefel, L., Rubenstein, R., Schwartz, A.E., 1999. Using Adjusted Performance Measures for Evaluating Resource Use. Public Budgeting and Finance, 19, 67--87.

Stiefel, L., Schwartz, A.E., Bel Hadj Amor, H., Kim, D.Y., 2005. Adjusted Measures of School Performance: A Cross-State Perspective, in: Stiefel, L., Schwartz, A.E., Rubenstein, R., Zabel, J. (Eds.), Measuring School Performance and Efficiency: Implications for Practice and Research, 2005 Yearbook of American Education Finance Association. Larchmont, NY: Eye on Education, 17--36.

Warren, J.R., Edwards, M.R., 2003. The Impact of High Stakes Graduation Tests on High School Diploma Acquisition. Paper prepared for presentation at the Fourth Annual Undergraduate Research Symposium at the University of Washington, May 2001.

Table 1. Descriptive statistics, student level, 2001 high school cohort

	All high school students			CUNY enrollees		
	(1) Mean	(2) Minimum	(3) Maximum	(4) Mean	(5) Minimum	(6) Maximum
Student is a high school graduate, in four years	0.62	0.00	1.00	1.00	1.00	1.00
Student applied to CUNY	0.42	0.00	1.00	1.00	1.00	1.00
Student enrolled in CUNY	0.26	0.00	1.00	1.00	1.00	1.00
First semester CUNY GPA	2.57	0.07	4.00	2.57	0.07	4.00
Verbal SAT CUNY Applicants[19]	439.76	200.00	800.00	430.55	200.00	800.00
Math SAT CUNY Applicants	471.97	200.00	800.00	461.18	200.00	800.00
English Regents' score	68.91	1.00	100.00	73.06	2.00	100.00
Sequential 1 Regents' score	67.55	1.00	100.00	68.90	22.00	99.00
Student is female	0.52	0.00	1.00	0.59	0.00	1.00
Age as of 2001	17.33	14.00	24.00	17.24	14.00	24.00
Student is white	0.18	0.00	1.00	0.23	0.00	1.00
Student is black	0.34	0.00	1.00	0.28	0.00	1.00
Student is Hispanic	0.32	0.00	1.00	0.28	0.00	1.00
Student is Asian	0.16	0.00	1.00	0.20	0.00	1.00
Student is native American	0.00	0.00	1.00	0.00	0.00	1.00
Student is eligible for free lunch	0.78	0.00	1.00	0.76	0.00	1.00
Student is eligible for reduced-price lunch	0.06	0.00	1.00	0.07	0.00	1.00
Student is not eligible for school lunch	0.15	0.00	1.00	0.17	0.00	1.00
Student is foreign-born	0.23	0.00	1.00	0.27	0.00	1.00
Student is native-born	0.77	0.00	1.00	0.73	0.00	1.00
Student entered the system before 9th grade	0.83	0.00	1.00	0.82	0.00	1.00

[19] SAT scores are reported for CUNY applicants only.

Table 1 (Continued). Descriptive statistics, student level, 2001 high school cohort

	(1) Mean	(2) Minimum	(3) Maximum	(4) Mean	(5) Minimum	(6) Maximum
Student entered the system in the 9th grade or later	0.09	0.00	1.00	0.09	0.00	1.00
Student is an English-language learner	0.12	0.00	1.00	0.13	0.00	1.00
Student speaks English at home	0.50	0.00	1.00	0.45	0.00	1.00
Student receives part-time education services	0.05	0.00	1.00	0.03	0.00	1.00
8th grade reading z-score	0.15	2.94	2.98	0.32	2.94	2.98
8th math reading z-score	0.16	2.65	2.16	0.35	2.51	2.16
LAB percentile (8th grade)	19.79	1.00	99.00	23.90	1.00	99.00

In columns (1), (2) and (3), N=50,494, except student applied to CUNY and student enrolled in CUNY (N=31,020), first semester CUNY GPA (N=11,718), verbal SAT (N=16,335), math SAT (N=16,336), English Regents' score (N=40,283), Sequential 1 Regents' score (N=39,952), race (N=50,402), eligibility for school lunch (N=34,995), time of entry into the system (N=46,152), 8th grade reading z-score (N=34,663), 8th grade math z-score (N=36,454), and 8th grade LAB percentile (N=6,083). See data section of paper for description of how sample is derived.

In columns (4), (5) and (6), N=13,342, except first semester CUNY GPA (N=11,718), verbal SAT (N=10,423), math SAT (N=10,424), English Regents' score (N=13,150), Sequential 1 Regents' score (N=12,726), race (N=13,340), eligibility for school lunch (N=9,290), time of entry into the system (N=12,020), 8th grade reading z-score (N=9,111), 8th grade math z-score (N=9,555), and 8th grade LAB percentile (N=1,612).

Table 2. Descriptive statistics, New York City high-school level, 2001

Variable	Mean	Minimum	Maximum
Enrollment (SBER)	1769	151	4631
Performance			
First semester GPA	2.52	1.44	3.33
Recentered SAT Verbal Score	419.61	300.71	677.52
Recentered SAT Math Score	440.85	339.29	720.59
English Regents' score	65.43	36.63	95.02
Sequential 1 Regents' score	65.96	55.57	88.83
% took the English Regents' exam	79.49	42.31	100.00
% took the Sequential 1 Regents' exam	79.03	19.05	100.00
% who graduated within four years	56.94	13.47	100.00
% of graduates who apply to CUNY	67.86	13.60	92.31
% of graduates who enroll in CUNY	42.00	6.40	65.77
Demographics			
% female	52.85	4.91	83.95
% white	13.20	0.00	79.55
% black	39.76	2.89	97.40
% Hispanic	35.25	2.56	92.27
% Asian	11.17	0.00	67.14
% Native American	0.35	0.00	2.94
% Free lunch eligible (8^{th} grade)	57.95	22.38	82.38
% Reduced price lunch eligible (8^{th} grade)	4.46	0.00	10.61
% Not eligible for lunch (8^{th} grade)	8.77	0.00	43.13
% foreign born	20.10	0.00	98.53
% native born	79.90	1.47	100.00
% who entered the system before 9th grade	84.23	1.10	100.00
% who entered the system in 9th grade or later	7.99	0.00	39.56
% English language learner	10.33	0.00	90.52
% in part-time special education (8^{th} grade)	5.00	0.26	11.76
CTB Z-Score (8^{th} grade)	0.03	-0.69	2.13
CAT Z-Score (8^{th} grade)	0.04	-0.66	2.07
% took ZCTB (8^{th} grade)	70.91	0.00	93.22
% took ZCAT (8^{th} grade)	74.11	0.00	94.59
Resources			
% teachers in this school for more than 2 years	70.94	24.20	100.00
% teachers with more than 5 years of teaching	60.76	14.80	89.50
% teachers fully licensed/permanently assigned	83.86	52.30	100.00
% teachers with a masters or higher	79.80	50.00	95.20

Table 2 (Continued). Descriptive statistics, New York City high-school level, 2001

Variable	Mean	Minimum	Maximum
Resources (continued)			
Total expenditure per pupil	10,171	7,308	39,409

School means come from aggregated student-level data unless otherwise indicated.

N=148, except eligibility for school lunch and 8th grade z-scores (N=147), part-time special education (N=143), teachers in the school for more than two years (N=135) and the other teacher variables (N=136).

Table 3. OLS Regression Results: High school performance of all high school students

	English Regents’	Graduated
Female	1.850***	0.083***
	(0.113)	(0.005)
Age as of cohort year	-5.036*	-0.688***
	(2.811)	(0.173)
Age square	0.094	0.017***
	(0.079)	(0.005)
Black	-2.156***	-0.025**
	(0.232)	(0.012)
Hisp	-1.773***	-0.077***
	(0.208)	(0.008)
Asian	-0.690***	0.016**
	(0.246)	(0.008)
Native American	-2.030**	-0.084**
	(0.824)	(0.035)
Free lunch 8th grade	-0.005	-0.012**
	(0.107)	(0.006)
Reduced lunch 8th grade	-0.112	-0.006
	(0.177)	(0.009)
Native born	1.744***	-0.065***
	(0.398)	(0.016)
Entered system before 9th grade	-0.670*	-0.077***
	(0.387)	(0.016)
Native born * entered system	-2.554***	-0.042***
before 9th grade	(0.383)	(0.015)
ELL	-6.187***	-0.088***
	(0.373)	(0.019)
Native born * ELL	2.432***	0.053**
	(0.506)	(0.021)
Speaks English at home	-0.386***	-0.041***
	(0.145)	(0.007)
PTSE 8th grade	-2.988***	-0.025**
	(0.262)	(0.011)
CTB Z-score 8th grade	4.636***	0.107***
	(0.119)	(0.005)
CTB Z-score 8th grade square	-0.231***	-0.023***
	(0.064)	(0.002)
CTB Z-score 8th grade cube	-0.177***	-0.005***
	(0.022)	(0.001)

Table 3 (Continued). OLS Regression Results: High school performance of high school students

	English Regents'	Graduated
CAT Z-score 8th grade	2.298***	0.145***
	(0.134)	(0.007)
CAT Z-score 8th grade square	-0.122	-0.021***
	(0.084)	(0.002)
CAT Z-score 8th grade cube	-0.237***	-0.016***
	(0.051)	(0.002)
LAB percentile 8th grade	0.030***	0.002***
	(0.005)	(0.000)
Constant	129.732***	6.889***
	(24.564)	(1.514)
Observations	40283	50494
R-squared	0.47	0.25
Fixed effects		
Mean	1.05e-08	-4.78e-11
Standard deviation	2.99	0.11
Joint F on fixed effects	26.76***	16.71***

Notes: *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. Robust standard errors are in parentheses. All models include missing dummies for free lunch, time of entry into the system, part-time special education, z-scores, and LAB percentile. The Regents' model also includes dummies for the year the Regents' were taken. The number of observations differs across models due to missing data on the Regents'. The graduation model was estimated with the sample of students for whom the Regents' are reported and the results were unchanged.

Table 4. Pearson correlations between high school outcomes for all students

Raw measures

	Graduation	English Regents'	Sequential I Regents'
Graduation	1	0.78 <.0001	0.75 <.0001
English Regents'		1	0.79 <.0001
Sequential I Regents'			1

Fixed effects

	Graduation	English Regents'	Sequential I Regents'
Graduation	1	0.47 <.0001	0.50 <.0001
English Regents'		1	0.49 <.0001
Sequential I Regents'			1

Note: Underlying regressions control for student characteristics

Table 5. OLS Regression Results: High school and college performance of enrolled CUNY students

	GPA	Verbal SAT	English Regents'
Female	0.200***	-2.428**	1.680***
	(0.015)	(1.209)	(0.140)
Age as of cohort year	-0.374	-78.709***	-2.919
	(0.251)	(27.912)	(3.372)
Age square	0.010	1.762**	0.050
	(0.007)	(0.770)	(0.095)
Black	-0.168***	-17.534***	-1.664***
	(0.027)	(3.178)	(0.259)
Hisp	-0.158***	-7.183**	-0.961***
	(0.022)	(2.846)	(0.219)
Asian	-0.050**	-12.928***	-0.823***
	(0.022)	(2.950)	(0.226)
Native American	-0.315*	-5.358	-1.961
	(0.188)	(16.040)	(1.276)
Free lunch 8th grade	-0.036	-0.437	0.010
	(0.028)	(2.145)	(0.179)
Reduced lunch 8th grade	-0.068*	0.252	-0.139
	(0.037)	(2.822)	(0.269)
Native born	-0.134***	35.411***	1.674***
	(0.048)	(5.350)	(0.461)
Entered system before 9th grade	-0.044	5.935	-0.308
	(0.050)	(5.132)	(0.436)
Native born * entered system	0.066	-37.138***	-2.225***
before 9th grade	(0.059)	(6.048)	(0.497)
ELL	-0.012	-67.926***	-4.602***
	(0.040)	(3.672)	(0.407)
Native born * ELL	0.117*	31.643***	2.347***
	(0.070)	(6.844)	(0.515)
Speaks English at home	-0.054***	6.965***	-0.003
	(0.018)	(1.962)	(0.161)
PTSE 8th grade	-0.071	-15.135***	-2.422***
	(0.045)	(5.775)	(0.385)
CTB Z-score 8th grade	0.130***	57.600***	3.959***
	(0.017)	(1.485)	(0.147)
CTB Z-score 8th grade square	0.044***	5.314***	0.062
	(0.010)	(1.188)	(0.103)
CTB Z-score 8th grade cube	-0.023***	-3.849***	-0.182***
	(0.004)	(0.510)	(0.046)

Table 5 (Continued). OLS Regression Results: High school and college performance of enrolled CUNY students

	GPA	Verbal SAT	English Regents'
CAT Z-score 8th grade	0.081***	23.373***	1.895***
	(0.021)	(2.151)	(0.152)
CAT Z-score 8th grade square	-0.010	1.628*	-0.036
	(0.011)	(0.871)	(0.072)
CAT Z-score 8th grade cube	0.001	-2.798***	-0.156***
	(0.008)	(0.652)	(0.040)
LAB percentile 8th grade	0.001	0.428***	0.033***
	(0.001)	(0.092)	(0.008)
Number of semesters between graduation and enrollment	0.128***		
	(0.016)		
Number of semesters between graduation and enrollment square	-0.011***		
	(0.002)		
Constant	6.424***	1,270.114***	112.014***
	(2.302)	(253.531)	(30.463)
Observations	11718	10423	13150
R-squared	0.10	0.56	0.39
Fixed effects			
Mean	2.64e-10	-3.40e-08	3.44e-09
Standard deviation	0.14	23.50	2.36
Joint F for fixed effects	2.12***	6.94***	9.94***

Notes: *, ** and *** indicate significance at the10%, 5% and1% levels respectively. Robust standard errors are in parentheses. All models include missing dummies for free lunch, time of entry into the system, part-time special education, z-scores, and LAB percentile. The Regents' model also includes dummies for the year the Regents' were taken. The number of observations differs across models due to missing data on the Regents' and SAT scores. Models were estimated with the sample of students for whom all three variables are reported (N=9,292) and the results were unchanged.

Table 6. Pearson correlations between the fixed effects from high school and college outcome models of CUNY enrollees

	English Regents’	Sequential I Regents’	Verbal SAT	Math SAT	GPA
English Regents’	1	0.35	0.54	0.50	0.25
		<.0001	<.0001	<.0001	0.00
Sequential I Regents’		1	0.23	0.43	0.03
			0.00	<.0001	0.75
Verbal SAT			1	0.80	0.47
				<.0001	<.0001
Math SAT				1	0.41
					<.0001
GPA					1

Note: Underlying regressions control for student characteristics

Table 7. Pearson correlations of fixed effects from outcomes models for high school students and CUNY enrollees

	Enrollees		
	English	Sequential I	GPA
High school students			
Graduation	0.23	0.29	-0.04
	0.01	0.00	0.61
English	0.89	0.37	0.24
	<.0001	<.0001	0.00
Sequential I	0.35	0.88	0.05
	<.0001	<.0001	0.51

Note: Underlying regressions control for student characteristics. Enrollees and high school students are different samples.

Table 8. OLS Regression Results: Model of college GPA controlling for high school success, CUNY enrollees

	GPA
Female	0.181***
	(0.015)
Age as of cohort year	-0.185
	(0.241)
Age square	0.005
	(0.007)
Black	-0.111***
	(0.026)
Hisp	-0.115***
	(0.022)
Asian	-0.068***
	(0.023)
Native American	-0.272
	(0.181)
Free lunch 8th grade	-0.038
	(0.028)
Reduced lunch 8th grade	-0.065*
	(0.036)
Native born	-0.154***
	(0.045)
Entered system before 9th grade	-0.026
	(0.044)
Native born * entered system	0.075
before 9th grade	(0.048)
ELL	0.112***
	(0.041)
Native born * ELL	0.085
	(0.068)
Speaks English at home	-0.049***
	(0.018)
PTSE 8th grade	0.017
	(0.045)
Number of semesters between	0.106***
graduation and enrollment	(0.016)
Number of semesters between	-0.008***
graduation and enrollment square	(0.002)
English Regents' score	0.012***
	(0.001)

Table 8 (Continued). OLS Regression Results: Model of college GPA controlling for high school success, CUNY enrollees

	GPA
Sequential I Regents' score	0.005***
	(0.001)
Verbal SAT	0.001***
	(0.000)
Math SAT	0.000***
	(0.000)
Constant	4.250*
	(2.192)
Observations	11718
R-squared	0.14
<u>Fixed effects</u>	
Mean	2.34e-10
Standard deviation	0.12
Joint F for fixed effects	1.84

Notes: *, ** and *** indicate significance at the10%, 5% and1% levels respectively. Robust standard errors are in parentheses. The model includes missing dummies for free lunch, time of entry into the system, part-time special education, and Regents' and SAT scores.[20]

[20] This model was estimated with the 8th grade scores included as well. Results were almost identical. The one significant difference is that being native-born and time of entry into the system have a joint effect on GPA. The correlation between the fixed effects for the two models is 0.9989.

Authors' Contact Information

Hella Bel Hadj Amor

Institute for Education and Social Policy

Steinhardt School of Education, New York University

82 Washington Square East, 7 fl

New York, N.Y. 10003

hella.bel@nyu.edu

Amy Ellen Schwartz

Institute for Education and Social Policy

Steinhardt School of Education, New York University

82 Washington Square East, 7 fl

New York, N.Y. 10003

amy.schwartz@nyu.edu

Leanna Stiefel

Robert F. Wagner Graduate School of Public Service

New York University

295 Lafayette Street, 2 fl, Room 3098

New York, N.Y. 10012

leanna.stiefel@nyu.edu