TEST SCORE GAPS IN NEW YORK STATE SCHOOLS: WHAT DO FOURTH AND EIGHTH GRADE RESULTS SHOW?

Leanna Stiefel and Amy Ellen Schwartz
Faculty
Wagner Graduate School

and

Colin Chellman
Research Associate
Institute for Education and Social Policy

New York University
September, 2003

Executive Summary

In this report, we analyze performance gaps by race/ethnicity, income and gender in New York State schools using fourth and eighth grade math and English language test results. While previous studies typically focus on gaps between subgroups at the district, state or national level, this study explores the school-level gaps in performance.

Our results highlight the legacy of residential racial segregation – many schools have too few whites or non-whites to allow a meaningful calculation of the subgroup test performance or test score 'gap' between groups within a school. Although our minimum subgroup size was six, which is relatively small, only 45.7% of elementary schools were found to have enough whites and non-whites to calculate meaningful gaps. About one-third of elementary and middle schools are predominantly white and one fifth predominantly non-white. While roughly one quarter of all fourth graders and one fifth of all eighth graders attend predominantly white schools, twenty-three percent of fourth graders and fifteen percent of eighth graders attend predominantly non-white schools. Thus, integrated schools, disproportionately located in New York City and the downstate suburban districts, educate only about half of the students in the State.

The implication is that the state-wide test score gaps significantly reflect both gaps between segregated schools and within integrated schools. Our results indicate that these gaps differ substantially – test score gaps between racially segregated schools are over 2.5 times greater than the gaps in racially/ethnically mixed schools.

Among integrated schools, race/ethnicity gaps are negatively related to the non-white performance in fourth and eighth grade ELA, and to the performance of blacks and Hispanics, in particular. Put simply, the white/non-white gap in performance is, generally, lower in schools with higher performance of non-whites, overall, and blacks and Hispanics, in particular**.**

While income segregation is less profound, well over one quarter of elementary schools and one fifth of middle schools are economically homogenous – their student body is either predominantly advantaged or disadvantaged. Again, the gap between segregated schools is considerably greater than the gap within mixed income schools – well over two times greater. Advantaged schools are disproportionately located in suburban districts, while the distribution of mixed schools is proportionate to the overall distribution of schools. Not surprisingly, predominantly disadvantaged ones, disproportionately located in New York City and the Big Four Cities, educate a poorer and less English-proficient student body.

New York State has few single sex schools, so gender gaps can be calculated in the vast majority. In the ELA assessments, gender gaps favor females in both grades; in math, the fourth grade performance is comparable, but in the eighth grade, males perform slightly better.

We compared the characteristics of schools that "beat the odds" to those with traditional gaps. Although the majority of schools exhibit traditional gaps, in all cases a significant number show comparable performance between subgroups and some show non-traditional gaps in which the relative performance of subgroups is reversed - the performance of non-whites, for example, exceeds the performance of whites. Interestingly, schools with comparable performance are distributed unevenly across the state and have somewhat different characteristics than those with traditional gaps. As an example, schools with small race gaps are disproportionately found in New York City and downstate suburbs, and show disproportionately large shares of Asian and small shares of black students.

Our analyses highlight the challenge faced in reporting school-level subgroup test scores and holding schools accountable for subgroup performance as conceived in the federal *No Child Left Behind* legislation. School policymakers face a dilemma – high minimum numbers of students in a subgroup, necessary for statistical reliability, will come at the 'cost' of coverage – fewer schools will be held accountable. At the same time, our analyses of integrated schools reveal a significant number of schools with comparable performance across races, genders, and income groups and, we find that smaller gaps are more likely to reflect higher performance by blacks and Hispanics, than low performance by whites.

I. Introduction

Research reports and headlines are rich with news concerning the test score gaps between blacks and whites (Jencks and Phillips, 1998), between males and females[1] (College Board, 2001), and between rich and poor[2] (Mayer, 2001; College Board, 2001). Whether these disparities are due to inadequate academic preparation (e.g., a paucity of pre-school reading at home, fewer math and science courses offered or taken by certain students), bias in the questions on standardized tests, negative psychological reactions to the testing environment, low teacher expectations, or yet other factors is hotly debated, while research efforts have been constrained by the dearth of data. Most studies are based on relatively small samples of students, but with the implementation of the "No Child Left Behind" legislation in 2002, all states receiving federal money will be required to report test scores for all schools, and, most importantly, to disaggregate those scores by various subgroups by school. New York State is ahead of most others, having released 2000-01 fourth and eighth grade math and English language test scores for every school in the state, in total and by subgroup.

In this report, we use New York State's fourth and eighth grade math and English language test results to describe and analyze the school-level gaps in performance between groups defined by race/ethnicity, income and gender.[3] In what follows, we will refer to schools with fourth grades as elementary schools and schools with eighth grades as middle schools, although there are 207 schools that serve both grades and are, therefore, included in both groups. As detailed below, our analysis uses data on 2262 schools serving fourth graders and 1074 schools serving eighth graders.

Note that use of the school as the unit of analysis in this type of study is relatively new. Studies investigating disparities in test scores typically focus on gaps measured at the district, state or national level.[4,5] While we know of no statistical studies that have examined the gaps in test scores at the school level, the reporting requirements of "No Child Left Behind" may quickly

---

[1] For example, national mean scores of college-bound seniors in 2001 are as follows: *Verbal* – Male (509), Female (502), and Total (506); *Math* – Male (533), Female (498), and Total (514). See College Board, 2001.

[2] National mean scores for the richest and poorest seniors taking the SAT in 2001: *Verbal* – Family Income < $10,000 (421) and Family Income > $100,000 (557); *Math* -- < $10,000 (443) and > $100,000 (569). Again, see College Board, 2001.

[3] Data on English language learners and on migrants are available, but too few schools have adequate numbers of students in more than one group to warrant analysis.

[4] Schwartz and Stiefel with colleagues Ellen and O'Regan have drafted a paper on test score gaps using New York City pupil and school data entitled "Beyond the Black-White Divide: Ethnicity, Segregation, and Academic Performance in a Multiethnic City" and are in the process of producing a revised version.

[5] Similar analyses have been produced in a limited number of other states that, like New York, have released student performance data disaggregated by sub-group. See, for example, Myers, 2001.

spur new research in that direction both because of the school-level subgroup test score data that will become available and because of the increased attention that will be focused on understanding the causes and consequences of disparities in subgroup performance within schools. Thus, we hope that this study will provide a foundation for future work in this area.

We begin with an examination of the distribution of test score gaps between race groups across the state. Are the familiar gaps at the aggregate level observed in most schools? Do whites always outperform non-whites or are there schools that "beat the odds" by producing parity in performance across race groups? To preview, this analysis highlights the continuing legacy of residential racial segregation – many schools have too few students in one racial subgroup or another to allow a meaningful calculation of a test score 'gap' between groups within a school. As described below, the implication is that the state-wide gap in test scores reflects both gaps between segregated schools and within integrated schools. Focusing, then, on the subset of integrated schools, we examine the schools that beat the odds and attempt to identify the ways in which they differ from others. In related analyses, we examine the correlation between the size of the test score gaps and other school (and district) characteristics, with an eye toward gaining insight into policies and/or practices that might ameliorate these test score gaps. By looking at schools that are beating the odds, we hope to begin to gain insight into the strategies that can be used to replicate this success.

We perform similar analyses of the disparity in test scores between economically advantaged and disadvantaged students and between females and males. These analyses are less complicated, in some measure, because of the greater prevalence of schools serving students with both of these dimensions.

We begin with a brief description of the sources for the data in section II and a summary of the statewide pass rates in section III. We then proceed to analyze gaps by race/ethnicity in section IV, by income in section V, and by gender in section VI. In section VII we conclude.

II. Data: A Brief Overview

Data on test scores, school characteristics, and district characteristics were provided by the New York State Education Department (SED) for academic year 2000-01 for all schools in the state. More specifically, we merged the following SED data sets:

- School Report Card data (SRC) for fourth and eighth graders, for the English Language Assessment (ELA) and the math assessment. These data contain the total number of students tested in each school and the number of students with scores at level 3 (meets standards) and

at level 4 (exceeds standards)[6]. In addition, these data are provided separately for each of five race/ethnicity groups, two income groups, and males and females, by school.[7]

- Institutional Master File (IMF) data for each school on selected school-level characteristics, such as enrollment and average percent of students registered for free lunch.
- State of Learning (Chapter 655) and Fiscal Profile data on selected district-level characteristics, such as expenditures per pupil and a needs/resource capacity index.

.

III. Statewide Pass Rates on Fourth and Eighth Grade English and Math Tests

Tables III.1 and III.2 show statewide pass rates for English and math tests. Note that a trivial number of schools (and their students) are excluded from these analyses either (1) because test data were available for five or fewer students, making statistical comparisons of scores extremely unreliable, or (2) because more than 40% of the students had Individual Education Plans (i.e., IEP's for disabled students), in which case the presence of a substantial special education population may confound interpretation. These exclusions make no difference to the patterns and make the subsequent school gap analyses more reliable.[8]

As shown in Table III.1, a much higher proportion of fourth graders than eighth graders pass each test, whites do better than nonwhites, and students advantaged by income do better than those who are disadvantaged. In fourth grade, females perform better than males in English but perform similarly to males in math. By eighth grade, the English advantage persists for females, and males perform slightly better in math.

Table III.2 breaks the non-white students into four, non-overlapping ethnic/racial groups. In both fourth and eighth grades, whites and Asians outperform Hispanic, black and American Indian students. In fourth grade, whites perform slightly better than Asians, but by eighth grade, this pattern is reversed.

These results are well-known and replicate the patterns reported in other studies. In this report, we go further to ask what is happening within schools in the state. Do the patterns of gaps obtain at the school level as well? Are there exceptions and, if so, what are the characteristics of schools that show negligible gaps?

---

[6] Performance level descriptions are taken from the glossary of statistics for public school districts in the description of Chapter 655 data, 2002 (School Year 2000-2001) <www.emsc.nysed.gov/irts/ch655_2002/glossary_June_2002.pdf>.

[7] Note that the data are cross-tabulated by each characteristic separately. We do not have data by subgroups jointly defined by race, income and sex.

[8] Appendix Tables III.1 (Race), Appendix III.2 (Gender), and Appendix III.3 (Income) show how we obtain the sample of schools with more than five students tested from the original data set of all schools. Tables III.4 shows the pass rates for all students (that is, including students in schools with five or fewer tested).

IV. School Test Score Gaps by Race/Ethnicity

Differences in academic performance by race are not new. For example, the National Assessment of Educational Progress (NAEP) shows such gaps in test scores dating back to the 1970's when NAEP was first administered.[9] In addition, performance indicators such as graduation rates and grade completion rates were vastly different in the early part of the century for blacks and whites. What has changed over the twentieth century is the degree to which children of different races attend the same schools -- more schools are integrated in 2000 than were integrated in 1950. The persistence of performance gaps is, then, particularly troubling to those who had hoped that integration would eliminate these disparities. Notice, however, that this increase in integration does not mean that segregated schools have disappeared. In fact, in New York State, a large fraction of elementary and middle schools educate a population of students predominantly white or predominantly non-white. The implication is that statewide gaps in test scores reflect both differences in scores *across* segregated schools and differences *within* integrated schools. Thus, we turn next to an analysis of racial diversity.

A. Racial Heterogeneity in New York State Schools

To understand the racial heterogeneity of New York State schools, Table IV.1 examines the distribution of schools (and the students who attend them) across three categories: predominantly white, non-white, or racially mixed. Schools are classified as predominantly white (non-white) if there are more than five white (non-white) students and five or fewer non-white (white) students. Mixed schools have more than five white students and more than five non-white students. In this and subsequent analyses, the sample includes only students tested. As noted above, our analyses exclude schools in which test data were available for five or fewer students and/or more than 40% of the students had Individual Education Plans.

Of the 2,262 elementary schools in our sample, we find that more than 34% are predominantly white and these schools educate around 26% of fourth graders. Twenty percent of schools (with 23% of fourth grade students) are predominantly non-white, and over 45% (with 51% of the fourth grade students) are integrated, educating a mix of white and non-white students. In the 1,074 middle schools, roughly a third are predominantly white (enrolling more than 18% of students), over 17% are predominantly non-white (over 14% of students), and 49% are mixed (nearly 67% of students). More than half of both elementary and middle schools are

[9] NAEP data also show that gaps between race groups are decreasing or at least remaining steady. See for example, Cook and Evans, 2000, and Jacobson et al, 2001.

racially homogeneous and these schools educate one-third to one-half of the students in the state. Thus, the need to look at test scores across segregated schools as well as within integrated schools is not merely of academic interest but rather is critical to understanding the academic performance of elementary and middle school students in New York State.

Table IV.2 shows further detail on the racial mix of students in each of the three racial heterogeneity categories, highlighting the substantial disparity in racial composition of these schools. In elementary schools, for example, while 61% of the fourth graders in the average school are white, the 775 predominantly white schools have virtually no non-white students (by construction). The 453 predominantly non-white schools are, on average, roughly 55% black, 37.8% Hispanic and 4.1% Asian. Interestingly, the composition of the mixed schools closely mirrors the composition of the 'average' school - roughly 60% of the fourth graders are white, 15% Hispanic, .5% are American Indian, 17-19.2% are black and 4.6%-7.7% are Asian. Further, the racial composition of the segregated schools, taken as a group, is not radically different from the composition of the mixed schools, thus offering the possibility that schools with more representative student bodies might be formed by integrating the students in the segregated schools only.

Note that while blacks, Hispanics, and American Indians show their highest representation in predominantly non-white schools, Asians are in their highest proportions in mixed schools.

B. Test Score Gaps and School Characteristics by Racial Heterogeneity of Schools

1. Elementary Schools. Table IV.3 displays test performance for elementary schools by racial heterogeneity category.[10] Most striking in this table is the difference in non-white pass rates between segregated and mixed schools.[11] While white performance is almost identical in segregated and mixed schools, non-white performance is much lower in the segregated than the mixed schools. On the ELA, just over 37% of non-whites in segregated schools pass while over 58% pass in mixed schools. A similar pattern obtains in math (45% pass in segregated schools and 68% pass in mixed schools.) The implication is that the disparity in performance between whites and non-whites is considerably higher within the set of segregated schools (i.e.,

[10] Note that the total pass rate and the racial pass rate for predominantly white and nonwhite schools differ slightly because of the small number of students of the other race group (fewer than 5 students) who attend the school. We use the rate for the specific race group in the analyses.

[11] This table and all others except where noted are based on school averages (not pupil-weighted averages). If pass rates differ by size of school, the two rates will not be the same.

predominantly white and predominantly nonwhite) than within the set of mixed schools.[12] On the ELA, the gap between whites and non-whites in segregated schools is 34.8 percentage points (71.9% – 37.1%) while the average gap within mixed schools is 13.5 percentage points; on the math test, the gap in segregated schools is 38.3 percentage points (83.3% – 45%) and within mixed schools it is 13.4 percentage points.

This result is important for a number of reasons. First, it shows that at the school level, a larger disparity exists *between* segregated schools than *within* mixed schools. While these data in no way imply a causal relationship between performance and segregation, the findings are intriguing, raising the hope that performance disparities might be ameliorated through integration.

Second, the results highlight the challenge posed by segregation that state and federal policy makers will confront in implementing the reforms mandated by *No Child Left Behind.* To be specific, schools will be required to report – and be held accountable for - test performance by racial subgroups only if there are adequate numbers of students in a group to allow for reliable statistics. Our results suggest that, even with a relatively low hurdle for the number of students tested, many schools in New York State will have subgroups too small to allow reporting of subgroup test scores. In our analysis, for example, we have required only that there are more than five students in the category 'white' or 'non-white' (a hurdle that is quite a bit lower than New York or most other states are likely to adopt) and found that only 45.7% of New York State's elementary schools would be required to report subgroup performance and calculate a test score gap.[13] Well over half of the schools will not be required to confront gaps because they are effectively or "statistically" segregated (meaning there are too few students to reliably measure gaps). Yet, the largest gap in performance is between segregated schools. Exempting segregated schools from accountability in this fashion may well be counter productive for two reasons. First, school-level accountability for subgroup test scores seems to provide little incentive or direction for ameliorating the large gaps in performance in segregated schools. Second, and no less important, it may create a disincentive to integrate - if doing so adds a new burden of accountability to the school. We return to these issues below.

---

[12] In the mixed schools the average gap in scores is the same as the gap in the average scores because the number of schools is the same; in the segregated schools, only the gap in the average scores can be calculated because the two groups do not attend the same schools.

[13] The lowest minimum subgroup sizes proposed for holding a school accountable are five in Maryland (and 95% confidence interval), 10 in Louisiana, South Dakota, and Utah (and a 99% confidence interval), 11 in New Hampshire (95% confidence interval), and 20 in New Jersey (and 95% confidence interval). New York State's current proposed minimum is 40 students (no confidence interval), although these figures are subject to change. See Lynn Olsen, "'Approved' Is Relative Term for Ed. Dept." *EdWeek*, August 6, 2003 <www.edweek.org/ew/ewstory.cfm?slug=43account.h22> and <www.edweek.org/ew/vol-22/43account.pdf> (August 6, 2003).

Having examined the pass rates in elementary schools and gaps in schools by racial heterogeneity, we now look at differences in district and school characteristics across the schools. Table IV.4 displays data on district characteristics by school racial heterogeneity. In terms of school location, we see striking differences between categories. Over 84% of the non-white schools (382) are in New York City, where less than 1% of predominantly white schools (about 5) are located in the City. There is little opportunity for these few white schools to provide effective integration for the non-white ones within the same jurisdiction. Slightly over 27.1% of mixed schools (280) are in New York City and perhaps the non-white schools could be melded more with these schools, but, in the end, with a white student enrollment of less than 16%, the City schools are not likely to become effectively integrated with white students.[14] White schools, on the other hand, are disproportionately located in rural areas and upstate suburbs (i.e., 26.1% of white schools are in rural locations, substantially exceeding the fraction of all schools located there at 10.2%; corresponding numbers for upstate suburbs are 45.3% compared to 22.7%). At the same time, mixed schools are disproportionately located in downstate suburbs, and, to a lesser extent in the Big 4 cities and upstate small cities. To some extent this pattern reflects the residential racial segregation that exists across New York State that may impede efforts to more fully integrate the schools.

The only geographical areas that seem to have opportunities for creating more integrated schools are upstate small cities and downstate suburbs, where there are significantly higher percentages of predominantly white than of predominantly non-white schools. The relatively large share of mixed schools in these same areas is, then, encouraging, if the existing success in creating integrated schools can be spread to schools nearby.

We turn next to examining the differences between white, non-white and mixed schools in the characteristics of their school districts. We use the need/resource capacity index developed and used by the New York State Education Department to classify districts by their ability to meet their expenditure needs out of local revenues. The code is based upon the Need/Resource capacity index, which uses the estimated percentage of K-6 poor children to capture relative need and the combined wealth ratio to capture revenue raising capacity.[15] The need/resource capacity index is meant to identify districts with similar needs for state or federal resources to effectively educate their students, distinguishing between urban, rural and suburban districts; between upstate and downstate suburbs and small cities, and between New York City and the four other

[14] This same pattern is even stronger in the Big-4 cities -- Buffalo, Rochester, Syracuse, and Yonkers -- where none of the schools is predominantly white.

[15] The combined wealth ratio is the ratio of district wealth (measured by weighting equally property value and income) per pupil to state average wealth per pupil.

big cities in New York State.[16] According to this index, for example, the Big 4 cities have the highest relative need (2.741), followed by New York City (1.775). And, there are significant numbers of High Need Rural and Urban-Suburban districts.

As shown in Table IV.4, predominantly white schools are disproportionately located in high need rural districts and average need districts, and dramatically under represented in New York City and the Big 4 districts.[17] Predominantly non-white schools, on the other hand, are overwhelmingly located in New York City (84.3%). Another 7.1 % of the non-white schools are located in Large City districts (i.e., the Big 4) and 5.5% in High Need Urban-Suburban districts. Given the relatively high need of New York City and the Other Large City districts, the implication is that only a small fraction of the non-white schools are located in Low – or even Average -- Need districts. The distribution of the mixed schools more closely mirrors the overall distribution of schools.

As shown in Table IV.4, white, non-white, and mixed schools differ in the racial composition of their districts, in a pattern consistent with expectations. On average, predominantly white schools are located in districts in which almost 95% of the students are also white. Interestingly, the average predominantly non-white school is located in a district in which only 84% of the students are non-white (15.9% of the students are white). Thus there seems to be greater opportunity for integrating the non-white schools with white students elsewhere in their districts than for integrating the white schools, whose districts contain, on average, very few non-white students. Finally, note that the average mixed school is located in a district in which more than half of the students are white, more than one fifth are black and both Hispanics and Asians are over-represented. Thus, the segregation of schools seems to reflect, to a significant extent, the segregation that exists between districts.

The last panel of Table IV.4 examines differences in spending, wealth and other characteristics. To begin, note that mixed schools are located in districts that spend the most per pupil and have the highest average income and combined wealth ratio. Interestingly, while non-white schools are located in districts that have, on average, the highest percentage of poor students and white schools are in districts that have the lowest percentage of poor students, white schools are located in districts that have the highest property values per pupil, but the lowest income per pupil. Consistent with our other measures, white schools are located in districts with

[16] Note that these categories, defined and provided by SED, are mutually exclusive and exhaustive. See *New York State of Learning*, June 2001, Table 3.1, page 68**.**

[17] The challenges of balancing and meeting the needs of rural and urban areas are faced by many states. See Molly A. Hunter, "State-By-State Status of School Finance Litigations." Campaign for Fiscal Equity's

the lowest average dropout rate, the highest average college going rate and attendance rate, and the smallest size. At the other end of the spectrum, non-white schools are located in districts with the highest average drop out rate, lowest attendance race and percentage going to college, and largest size. Mixed schools, again, are distributed in such a way that the averages closely match the state overall. As noted previously for racial mix of students, the implication is that, taken as a group, the segregated schools are not so different from the mixed schools – the big differences are between the segregated-white schools and the segregated-non-white schools.

Turning next to school characteristics, in Table IV.5, we see stark differences between nonwhite and white schools. Predominantly non-white schools are largest and have the highest poverty and English language learner rates.[18] While there are very few whites in these schools, those who are there score very low on the tests (with average pass rates of 40.5% and 48.3% on the ELA and math tests respectively). In predominantly white schools, on the other hand, the few non-whites score particularly well. It is important to stress, however, that the small numbers of students in these subgroups make these results suggestive, at best.

In sum, there is considerable segregation in elementary schools, and test results differ across schools with different racial mixes as do several school characteristics.

2. Middle Schools. Results for middle schools are, for the most part, similar to those for elementary schools. Therefore, we discuss highlights and differences only, although we replicate all of the elementary school tables for middle schools.

Table IV.6 displays the pass rates by racial heterogeneity category of middle schools. Again, the largest differences in scores are between whites in white schools and non-whites in non-white schools. The gap in ELA is 26.7 percentage points versus 14.9 in mixed schools; the gap for math is 26.5 percentage points versus 16.9 in mixed schools.

Table IV.7 shows district characteristics for middle schools. Predominantly white schools are even more highly concentrated in rural areas (41.7%) than elementary schools and they are equally concentrated in upstate suburbs as elementary schools. In addition to New York City and the Big 4 cities, mixed schools are disproportionately located in upstate small cities and downstate suburbs. Only downstate suburbs have much opportunity to further mix white and non-white schools since all other areas have very small 'other' race/ethnicity populations.

Advocacy Center for Children's Educational Success with Standards. www.accessednetwork.org/litigation/TableofLitigationStatusfor50states.pdf> (July 17, 2003).

[18] These schools do not have the highest special education rates because they are primarily located in New York City, which educates a larger proportion of its special education students in schools with over 40% of special education students and these latter schools are removed from the data.

The needs/resource capacity index again shows that disproportionately white schools are located in high need rural areas and average need areas while non-white schools are found even more disproportionately in New York City (89.8%) than elementary schools.

Other relationships with enrollment and fiscal characteristics are similar to elementary schools and characteristics of schools in Table IV.8 show similar patterns to elementary schools. Having described the overall patterns, we next turn to an examination of schools that show test performance parity between race groups or even show higher performance among the non-whites.

C. Examining the Race Gap in Test Scores at the School Level

1. Description and Characteristics of Schools by Gap Category. Are any schools in the state "beating the odds" by producing pass rates that are similar for whites and non-whites or perhaps even reversed, with non-whites performing better? To answer this question, we turn to the racially mixed schools (1034 elementary schools and 527 middle schools), which are the only schools where such a question is relevant. After identifying schools that are beating the odds in this section, we then analyze their characteristics, including their composition and the performance by specific non-white groups (Hispanic, black, Asian and American Indian) in Section D. The purpose of these analyses is to show that it is possible for schools to have small gaps in performance across subgroups and to identify characteristics that are particularly associated with such "successful" schools. The hope is that this will provide the foundation for a search for policies and strategies that can be replicated and applied elsewhere.

To be specific, then, we classify schools based on their gaps in *both* ELA and math tests into five non-overlapping categories, as shown in the rows of Table IV.9. Schools with *traditional* (or typical) gaps are schools in which the white pass rate on both tests is more than five percentage points higher than the non-white rate. Schools with small gaps (or in which performance is comparable in the two groups) are defined as those in which the white pass rate is within five percentage points of the non-white pass rate (either above or below).[19] Schools with *non-traditional* pass rates are those in which non-white students pass at higher rates (over five percentage points higher) than whites. Of course, not all schools fall neatly into these categories for both of the state tests. Schools we term *non-traditional/small* are defined as those in which the pass rates on the ELA and math tests each fall into one of these different groups, and a final category *other* (or conflicting) are schools in which the relationship between the white and non-

white performance differs substantively between the two tests, either one traditional and one non-traditional or one traditional and one small gap on the two tests.

As shown in Table IV.9, the vast majority of elementary and middle schools exhibit traditional gaps (58% of elementary schools and 68% of middle schools). That is, white students significantly outperform non-white students by more than five percentage points. On the other hand, a sizeable number of schools show small gaps (7% and 6% respectively of elementary and middle schools). Further, in some schools non-whites outperform whites by more than five percentage points. Taken together, 18% of elementary schools and 14% of middle schools obtain higher or comparable performance for non-whites than whites. Regardless of whether disparate performance across race groups is desirable if it favors minorities, it is clear that there are significant numbers of schools that beat the odds (are in non-traditional or small gap categories), offering potential for policy guidance concerning the Board of Regents priority for closing the student achievement gap.

Table IV.10 shows the location of schools by gap classification. Elementary schools that beat the odds are disproportionately located in New York City and in the downstate suburbs. For example, together, these two locations account for over 85% of the 77 small gap schools. Again, in eighth grade, New York City and the downstate suburbs disproportionately beat the odds, and for this grade, the rural schools join them. These three locations account for over 90% of the 32 middle schools that have small gaps. Conversely, the Big 4 cities and the upstate small city schools are disproportionately represented in the category of schools with traditional gaps.

Are there school characteristics that differ across schools classified by the nature of test score gaps? Or, is there any evidence that schools that beat the odds differ from others? Table IV.11 is constructed to shed light on these questions and a few observations are in order. To begin, non-traditional gap schools are smaller for both grades, while in eighth grade, small gap schools are larger, on average. In the elementary schools, there are many characteristics in addition to smaller school enrollment that distinguish schools with gaps that beat the odds from other schools. Non-traditional gap schools are disproportionately poor and non-white – with relatively high shares of Asian and Hispanic students. They educate a larger share of children in special education and ELL programs. At the other end of the spectrum, elementary schools with traditional gaps have lower percentages of poor students and, again, higher percentages of students who are non-white - in this case, particularly black students.

Fewer characteristics differ across gaps for middle schools. In contrast to fourth grade, however, middle schools show higher percentages of English language learners in small and

[19] Unless this comparability is achieved by lowering both groups' scores.

small/non-traditional gap schools and higher percentages of students who are white. As before, middle schools with traditional gaps enroll a higher percentage of black students.

The striking similarities in the results for elementary and middle schools are the high percentages of Asians in schools that beat the odds and the higher percentage of blacks in schools with traditional gaps. These naturally lead to questions about the pass rates for subgroups of nonwhites. That is, when the white/non-white gap is *not* traditional, is this due to the performance of one - or perhaps two - race groups or do all race groups perform better? The next set of tables, IV.12 though IV.15b address this question.

2. Pass Rates and Gap Categories by Racial/Ethnic Subgroups. Beginning with elementary schools, Table IV.12 shows the subgroup pass rates for predominantly non-white and for mixed schools. It also shows the white pass rate compared to each racial subgroup's pass rate for mixed schools. For all subgroups, the pass rate in predominantly non-white schools is markedly lower than in mixed schools. This is the same result we found earlier for non-whites as a whole, but here we see that it holds consistently for every racial/ethnic group.

Note, however, that test score gaps in mixed schools can only be calculated for that set of the mixed schools that have at least one student test score reported for the specific racial/ethnic group and, thus, the number of schools for each subgroup is less than the total of mixed schools (1034). Here, we have not restricted the analyses to schools with minimum numbers of student test scores in each subgroup. Thus, a gap for a particular school and subgroup may reflect the performance of as little as one student in a subgroup, although the typical number is considerably higher. That said, the gaps show a consistent pattern for ELA and math tests: gaps are largest for black students, non-existent to negative (or non-traditional) for Asian students, and in-between (but always positive/traditional) for Hispanic and American Indian students.

The subgroup analyses of test score gaps, then, suggest an important question: are the white/non-white test score gaps driven by the performance of just one race/ethnicity group? More specifically, do schools that beat the odds achieve this distinction due to the performance of Asians alone, with other groups performing similarly to their performance in schools with the traditional gaps? The answer to this question, one way or the other, will suggest different policy interventions.

D. Subgroup Gaps

Tables IV.13a and b show the relationship between subgroup gaps on the ELA and math tests, respectively. The white/non-white gap category is presented across the top of the table (the

columns), while the rows present the classifications for each racial/ethnic group. To begin, note that, because not every school contains students of each race/ethnicity, there is a row labeled "no students" to show how many schools have none of the specific race/ethnicity. For Hispanics, for example, 55 schools of a total 1034 that are mixed racially have no Hispanic students. The next three rows are then a cross tabulation between the overall white/nonwhite test score gap and the white/Hispanic test score gap, with the numbers and percents calculated from the total of 1034 mixed race schools. Data for race groups are arranged similarly.

There are a number of ways to investigate whether specific race/ethnicity gaps accord reasonably well with the white/non-white gaps, but one of the best is to calculate the percent of times that the "worst case scenario" occurs – when the results conflict . Arguably, the worst cases are when the white/non-white gap is non-traditional or small while the specific race/ethnicity gap is traditional. That is, the school overall successfully ameliorates or reverses gaps, but an individual subgroup among the non-whites performs markedly worse than whites. We can calculate this percent from the table by dividing the number of schools with traditional subgroup gaps by the number of schools (with students of that subgroup) that have non-traditional or small overall white/nonwhite gaps. For Hispanics on the ELA fourth grade test (Table IV.13a), this rate would be calculated as: [(14 + 69)/(1034-55)] times 100 or 8.5%. The interpretation is that 8.5% of the mixed schools with at least one Hispanic student post non-traditional or small white/non-white gaps but poor performance of their Hispanic students. Comparable numbers for the other subgroups are: blacks 10.1%, Asians 3.8%, and American Indians 12.7%. These low numbers suggest that, in general, schools with a white/nonwhite gap that is non-traditional or small are unlikely to be associated with a subgroup rate that is traditional.

As shown in Table IV.13b, the rates for mathematics are even lower: Hispanics 8.0%, blacks 8.6%, Asians 1.8%, and American Indians 4.6%.

The subgroup gaps for middle schools are shown in Table IV.14. As in the elementary school analyses, all subgroup scores for both ELA and math are lower in predominantly non-white than in mixed schools. Again, black gaps are largest and Asian gaps are smallest. Here, Asian gaps are both non-traditional and quite high (-8.4 percent and -11.0 percent, in ELA and math, respectively).

As for the prevalence of traditional subgroup gaps in schools in which the white/non-nonwhite overall gap is non-traditional or small, the rates are again low, although not as low as in elementary schools. For ELA (math) the rates are: Hispanics 11.8% (14.8%), blacks 10.3% (12.9%), Asians 1.5% (1.1%), and American Indians 8.8% (9.9%).

Thus in terms of subgroups, we see that Asians do best, and blacks do worst, and there is a match between the high performing white/non-white gap schools and the subgroup gaps.

E. The White/Non-White Test Score Gaps and School and District Characteristics

In this final section, we use correlations and bivariate regressions to study the relationship between school and district characteristics and the sign and size of white/non-white test score gaps. Previous sections have examined the means of these characteristics across various categories of schools (by racial composition or gap classification), but to this point, we have not examined the extent to which the *size* of the gap systematically varies with these characteristics within as well as between the categories we used earlier. As in the last section, these analyses are best viewed as descriptive, rather than isolating causal relationships.

The graphs, beginning with Graph IV.1, show a scatterplot of the school-level data for New York State as well as the ordinary least squares (OLS) regression line, with the gap on the Y axis. The equation for the line, its R^2 or coefficient of variation, and the r or Pearson correlation coefficient are also displayed. All of the regressions are bivariate, capturing the relationship between two variables only, and as such they do not hold constant (or control for) the effect of any other variables that may also be related to the gap. Therefore, caution is warranted in interpreting the results (and drawing conclusions about causality) because the 'true' relationship between these variables is, undoubtedly, more complicated and there are likely to be other independent variables correlated with the one analyzed that are absent from the regressions. Thus, it is possible that the relationship portrayed is misleading - the independent variable we analyze could well be and likely is correlated with other independent variables, whose effects are being represented. Further, we present simple linear (straight line) relationships, while more complex nonlinear ones may better represent the pattern of the data.[20] Nevertheless, these analyses are valuable because they help to distinguish those variables important to investigate in more sophisticated analyses from those less likely to be important. Further, they are interesting in their own right because provide insight into relationships about which there has, traditionally, been a good deal of opinion but little evidence.[21]

We begin with elementary schools and concentrate our discussion on the ELA results for two reasons. First, many individuals consider proficiency in reading and English skills

[20] Quadratic relationships do often "fit" the data better, but only marginally so.

[21] If there is no (significant) bivariate relationship, there might still be a relationship in a multivariate equation. But, this situation is unlikely since all biases in the coefficient due to the effects of omitted variables would have to just equal each other, positive and negative. On the other hand, a positive or

prerequisite to other learning, and therefore a particularly crucial skill. Second, the results for the ELA test score gaps are qualitatively similar for math. Appendix tables contain the math results.

1. Elementary Schools. Begin, in Graph IV.1, with the relationship between the ELA and math gaps for elementary schools and observe that the two are positively related, with an R^2 of .35, meaning that 35% of the variation in the ELA gap is explained by variation in the math gap. Another way to understand the relationship is to note that the correlation coefficient is .59, which is moderately high and is the highest correlation we find in any of our analyses, below.[22] Notice that while a high correlation in the level of test performance is familiar, the finding that the gaps are correlated across tests is somewhat new, but explains why our results for the ELA and math are so similar, as noted above.

A nagging question in the gap analyses is whether a reduced gap reflects low test scores by one or both of the race/ethnicity groups. As an example, are gaps lower in schools with low scores for whites, rather than high scores for non-whites? Or are scores low for both groups? The next series of graphs addresses these questions. We begin by considering whether, or in what way, gaps are associated with overall performance. As shown in Graph IV.2, the relationship between the gap and the overall fourth grade pass rate is non-traditional, although explanatory power is low – the R^2 is less than 3%. This indicates that the gaps are not smaller in schools with lower overall performance, which is good news.[23]

Is there a systematic relationship between the gaps and performance of a particular race group? As shown in Graph IV.3, the relationship between the size of the gap and the percent of white fourth graders who pass is mildly positive, but, as before, the explanatory power is low (R^2 of .04). The regression coefficient indicates that an increase of 10 percentage points in the white pass rate is associated, on average, with in an increase in the ELA gap of 1.9 percentage points. Put differently, when white pass rates *decline*, the gap *declines*, but the magnitude and strength of the relationship are very small.

On the other hand, as shown in Graph IV.4, *increases* in the non-white fourth grade pass rate are associated with large *reductions* in the gap. Here, 35% of the variation in the gap is accounted for by the non-white pass rate and an increase in the non-white pass rate of 10 percentage points is associated across schools with a reduction of the gap, on average, by 4.5

---

negative relationship might change size or even direction when other variables are added, but it is not likely to disappear all together.

[22] In bivariate regressions, but not multivariate ones, the square of the correlation coefficient (r) is equal to the coefficient of determination (R^2). (.59 times .59 = .3481)

percentage points. It would seem that schools with higher non-white pass rates have lower gaps, offering the hope of a "win-win" situation in which disparities are ameliorated by improving the outcomes of low performers rather than reducing outcomes by high performers.

Does this negative relationship hold for all races and ethnicities? The next series of graphs addresses this question and the answer is "yes," although not to the same extent for each subgroup. For blacks and Hispanics (shown in Graphs IV.5 and IV.6), there is a somewhat strong negative relationship (R^2's of .15 and .14, respectively), but for Asians and American Indians (Graphs IV.7 and IV.8), the relationship is less strong (R^2 of .08 or 02), respectively. Thus, we find no evidence that high subgroup performance is systematically associated with larger gaps.

How does school size relate to the gap? Graph IV.9 suggests that enrollment is, on average, unrelated to the size of the gap. The R^2 is an incredibly small .004, which is not statistically different from zero. Does the same pattern obtain for different race groups? In a series of graphs shown in the appendix (Appendix Graphs IV.8-13), the racial composition of school enrollment is also seen to be unrelated to the size of the gap.

Two other school characteristics, percentage of students in poverty and percentage of students who are English Language learners, are often thought to make education more challenging and thus might be expected to show a positive relationship with the size of the test score gap. Graphs IV.10 and IV.11 indicate, however, no statistical relationship between these attributes at the school level and gaps in elementary schools.

We next turn our attention to district-level characteristics that might plausibly be correlated with ELA gaps in elementary schools. We analyze four of these: expenditures per pupil, the combined wealth ratio, the district percent of free and reduced lunch pupils, and district size as measured by enrollment. These relationships are displayed in Graphs IV.12 through 15. None of these variables is systematically related to the gap.

To summarize, what do these reveal? These simple correlations suggest only one set of variables is systematically related to the size of the fourth grade ELA gap - the pass rate of the subgroups. In particular, increases in the non-white pass rate as a whole, and increases in the black and Hispanic pass rates in particular, have moderately strong and negative associations with the overall white/non-white gap. As stressed earlier, while these correlations are merely suggestive, and additional work is clearly warranted to separate out the specific effects of each of these various potential determinants, these might well be regarded as encouraging – reducing the gaps in performance between whites and other race groups may well be accomplished by

[23] When R^2's are very low (<.01), we report that the variables are unrelated. Occasionally, such R^2's are "statistically" different from zero, but, in our view, these have little substantive import or policy relevance.

improving test scores of the other race groups.

2. Middle Schools. The relationships for middle schools (shown in graphs IV.16 through IV.30) are similar, for the most part, to those observed for elementary schools. In the interest of brevity, we comment only on those that show some difference.

The relationship between the ELA and math test score gaps in eighth grade are stronger than in fourth (Graph IV.16), as shown by the larger R^2 of .44 (compared to .35 for elementary schools) or a correlation coefficient of .66 (compared to .59). Likewise the relationship with the white pass rate is slightly stronger and larger in magnitude and the relationship with the non-white pass rate is slightly less strong and smaller (see Graphs IV.18 and IV.19). The subgroup pass rates are also slightly less strongly related to the gap than they were in elementary schools. All other relationships repeat those of elementary schools and are insignificant.

V. School Test Score Gaps by Economic Status (Income) of Students

Students are identified as economically disadvantaged (by income) if they are enrolled in free- or reduced-price lunch programs and advantaged if they are not.[24] As shown in Table V.1 (which replicates Table III.1), the statewide pass rate on both the ELA and math tests, in both elementary and middle schools, is roughly 30 percentage points higher for advantaged than disadvantaged students. As in our analysis of test score gaps across races, we begin with an exploration of the extent to which students from different economic backgrounds share schools in New York State (Section A). This is followed by an analysis of gaps and characteristics of schools by income heterogeneity (Section B). We then identify mixed income schools whose gaps are low or reversed and examine their characteristics (Section C). Again, we end with a regression analysis of the relationships between the size of the school-level gap and school and district characteristics (Section D).

A. Income Mix in New York State Schools

Following the same procedure as in the race analyses, we classify schools as advantaged (disadvantaged) if there are more than five students who are economically advantaged (disadvantaged) and five or fewer are disadvantaged (advantaged). Mixed schools have more than five students who are advantaged and more than five who are disadvantaged. Table V.2

[24] In New York City, free lunch eligibility is defined as family income less than or equal to 130% of the federal poverty line and reduced-lunch eligibility is defined as less than or equal to 185% of the poverty line.

presents the distribution of schools (and the students who attend) across income mix categories.

Of the 2,262 New York State elementary schools, over 16% are predominantly advantaged and these schools educate around 13% of fourth graders. Around 12% of these schools (with 11% of fourth grade students) are predominantly disadvantaged, and 72% (with 76% of the fourth grade students) educate a mix of advantaged and disadvantaged students. In the 1,074 middle schools, over 13% are predominantly advantaged (nearly 10% of students), nearly 8% are predominantly disadvantaged (nearly 5% of students), and nearly 79% are mixed (86% of students). Thus, a substantial share of the schools in the state is income segregated – that is, their students (representing a significant fraction of the whole) are predominantly advantaged or disadvantaged. The need to look at test scores across these income segregated schools as well as within income integrated or mixed schools is, thus, relevant, although not quite so urgent as for the race/ethnicity analysis, where the segregation was more profound. (See Table IV.1)

B. Test Score Gaps, School and District Characteristics and Economic Integration

1. Elementary Schools. Table V.3 displays fourth grade test pass rates by income mix category of schools.[25] Most striking in this table are the differences in advantaged and disadvantaged pass rates between income segregated and mixed schools.[26] The pass rates in both math and reading in segregated schools are higher for advantaged students and lower for disadvantaged students than comparable rates in mixed schools. On the ELA test, over 83% of advantaged students in advantaged schools pass while only about 70% pass in mixed schools. A similar pattern obtains in math (90% pass in advantaged schools and 78% pass in mixed schools.) For disadvantaged students, 34.8% pass the ELA test in disadvantaged schools while 51.6% pass in mixed schools. Once again, in math, the pattern is similar (43.4% pass in disadvantaged schools and 64.6% pass in mixed schools). As we saw earlier in the race/ethnicity analysis, the gap between groups is larger among the segregated schools (48.6% for ELA and 47.0% for math) than within mixed ones (18.1% for ELA and 13.7% for math), although the numbers of schools and students affected is much lower.[27]

---

[25] Note that the total pass rate and the pass rate for predominantly advantaged and predominantly disadvantaged schools differ slightly because of the small number of students of the other income group (fewer than 5 students) who attend the school. We use the rate for the specific income group in the analyses. We do not present the table showing these numbers, but it is available in the appendix and is comparable except for income to Table IV.2 for race/ethnicity.

[26] This table and all others except where noted are based on school averages (not pupil-weighted averages). If pass rates differ by size of school, the two rates will not be the same.

[27] This result differs from that of Section IV, where the white pass rates were similar in predominantly white and mixed schools, but the non-white rates were lower in predominantly non-white than in mixed schools.

Having identified the fourth grade pass rates and gaps in schools by income mix, we now look at differences in district and school characteristics across the schools. Table V.4 displays data on district characteristics. To begin, notice that almost 93% of the predominantly disadvantaged schools are located in New York State's five largest cities, with over 81% in New York City alone. Conversely, almost 86% of the predominantly advantaged schools are in the suburbs, either upstate (58%) or downstate (28%). The needs/resource capacity index is less informative here because of the necessarily close relationship between advantaged/disadvantaged status and the index. That said, the mixed schools, whose classification by index is less predictable than that of advantaged schools, are somewhat disproportionately high need or average need on the index. The relationship by race/ethnicity of district enrollment shows that predominantly advantaged schools are located in districts with average white enrollment of almost 88% while predominantly disadvantaged schools are located in districts with over 80% non-white enrollment. These data confirm the familiar positive correlations found between income and race across many different units (individuals, school districts, schools etc.).

Table V.5 displays school characteristics by income mix of elementary schools. Advantaged schools have the lowest average enrollment and disadvantaged schools the highest. Demographic characteristics follow the familiar patterns associated with income: advantaged schools have the lowest rates of students in poverty, in special education and who are English language learners; the disadvantaged schools have the highest rates, and the mixed schools are in-between. Because there are some students who are disadvantaged in the advantaged schools (although five or fewer by construction) and vice versa, it is possible to calculate a pass rate for each group across schools. In the advantaged schools, pass rates for both groups are higher than in any other group of schools. Likewise in disadvantaged schools the pass rates for both groups are lower than in any other group of schools. This result mirrors our findings for racially segregated schools, in which both groups did particularly well in predominantly white schools, particularly badly in predominantly non-white schools, and in-between in mixed schools (see Table IV.5). Again, these results are suggestive, at best, because of the small numbers of students in these "other" groups.

2. Middle Schools. Results for middle schools are, for the most part, similar to those for elementary schools. Therefore, we present highlights and differences only, although all tables that were presented for elementary schools are replicated for middle schools.

Table V.6 displays the pass rates by income mix of schools. Again the largest differences in scores are between advantaged students in advantaged schools and disadvantaged students in

disadvantaged schools. The gap in ELA is 35.1 percentage points, substantially larger than the 19.4 in mixed schools; the gap for math is 44.2 percentage points which is, again, larger than the 17.0 in mixed schools.

We now look at differences in district and school characteristics across the schools. Table V.7 displays data on district characteristics. In terms of school location, 94% of predominantly disadvantaged schools are located in New York State's five largest cities, with over 82% in New York City alone. Conversely, almost 78% of the predominantly advantaged schools are in the suburbs, either upstate (34%) or downstate (44%). The relationship by race/ethnicity of district enrollment is similar to that for elementary schools. Predominantly advantaged schools are located in districts with average white enrollment of almost 86% while predominantly disadvantaged schools are located in districts with over 80% non-white enrollment.

Table V.8 displays school characteristics by income mix. Again, advantaged schools are smallest, but here, mixed schools, rather than disadvantaged schools, are the largest. Again, demographic characteristics follow a familiar pattern: advantaged schools have the lowest rates of students in poverty and who are English language learners, disadvantaged schools have the highest rates, and mixed schools are in-between. In middle schools, in contrast to the elementary schools, the predominantly disadvantaged schools have the lowest share of special education students. Finally, in the advantaged schools, pass rates for advantaged students are higher than in any other group of schools, but for the disadvantaged, the rate is similar to that in mixed schools and higher than in disadvantaged schools. In disadvantaged schools the pass rates for both groups are lower than in any other group of schools.

C. Examining the Income Gap in Test Scores at the School Level

As before, we now consider whether there are any schools that are beating the odds by producing small gaps for disadvantaged and advantaged students (equal to or within a five percentage points). Schools that exhibit non-traditional gaps have pass rates for disadvantaged students that are over five percentage points higher than for advantaged students and it is not clear that this is desirable, although it is certainly less common than schools with traditional pass rates. The distribution of schools by grade and gap category is shown in Table V.9. Unlike the racial gap analysis, for income (and gender in the next part of the report), we separate the results for the ELA and the math tests.

A substantial fraction of the 1,629 elementary schools have small gaps in test scores: 13.8% and 18.9% for the ELA and for the Math tests respectively. An additional 6.8% and 9.9%

have gaps that are non-traditional. That leaves the overwhelming majority, between 70% and 80% of schools, with traditional gaps. Slightly lower percentages of middle schools have small gaps (11.9% for ELA and 17% for math) and slightly lower percentages have gaps that beat the odds (7.1% and 6.5%). Nevertheless, there are enough schools with small or non-traditional gaps in both grades to warrant an analysis to see if their average characteristics differ from schools with traditional gaps.

Table V.10 shows the location of schools by gap classification. Elementary schools that have non-traditional or small gaps, in both ELA and math, are disproportionately located in New York City and downstate *small cities*. In the racial analysis, New York City also exhibited favorable results, as did downstate *suburbs*. The downstate small cities replace the downstate suburbs in the income gap analysis.

Data in Table V.11 shed light on whether there are differences in average school characteristics across gap classifications for elementary schools. The answer is yes –traditional gap schools differ on a number of variables compared to small and non-traditional gap schools. These schools with traditional gaps are smaller, have lower poverty and English language learner rates, have lower percentages of students who are disadvantaged economically, and have higher advantaged and lower disadvantaged pass rates. In other words, these schools generally educate a less challenged group of students, do very well by their advantaged students, and do least well by their disadvantaged students.

Table V.12 displays the geographic location of middle schools by gap, indicating that schools with non-traditional and small gaps are disproportionately located in New York City and the other four big cities of the state, while schools with traditional gaps are disproportionately in rural areas and in downstate and upstate suburbs. This result holds for both ELA and math test gaps. New York City houses a disproportionate share of the schools that are non-traditional on both tests for schools with both fourth and eighth grades.

Table V.13 enumerates school characteristics by gap classification for middle schools and results are similar to elementary schools - schools with traditional gaps have fewer students in poverty, English language learners, and students who are advantaged. In addition, similar to elementary schools, schools with traditional gaps produce higher advantaged and lower disadvantaged scores than the schools with small or non-traditional gaps.

In conclusion, New York State schools are somewhat "segregated" by income, although a higher percentage of schools include students from both income categories than include schools from both race categories. As with race, the gap in test scores across segregated schools is higher than within mixed schools and the location of segregated and mixed schools differs, with New

York City and the other four big cities housing a disproportionate share of disadvantaged schools as well as the downstate and upstate suburbs housing a disproportionate share of advantaged schools. The disadvantaged schools are disproportionately non-white and, of course, poor, on the basis of both school and district fiscal characteristics. Finally, about 20% to 30% of schools exhibit non-typical test score gaps and these schools are disproportionately located in New York City. The schools with traditional gaps educate less challenged students and they produce higher advantaged and lower disadvantaged test scores than the other groups of schools.

D. Income Gaps and District and School Characteristics

In this section, we follow the presentation on race/ethnicity and use correlations and bivariate regressions to study the relationship between district and school characteristics and the sign and size of advantaged/disadvantaged test score gaps. We again show a series of graphs of the best fitting line through the data, with the gap on the Y or vertical axis and the characteristic on the X or horizontal axis. The equation for the line, its R^2 or coefficient of variation and the r or Pearson correlation coefficient are also displayed. Again, when R^2's are low (below .01), we report that there is no relationship.

1. Elementary School Relationships. First, Graph V.1 shows the relationship between the ELA and math gaps for elementary schools, revealing that the two are positively related, with an R^2 of .32 and a correlation coefficient of .57. This is similar to the relationship we observed between the ELA math gaps in the race group analysis.

As before, we next explore the relationship between the gap and the subgroup test scores. Are the gaps smaller in schools with lower scores for advantaged students or do differences reflect higher performance by disadvantaged students? To answer these questions, graph V.2 plots the relationship between the gap and the overall fourth grade pass rate. Interestingly, in this case, the relationship is weak – little of the variation is explained (1%). Graph V.3 shows a positive and significant relationship between the size of the gap and the pass rate for advantaged fourth (R^2 of .20). The regression coefficient indicates that a decrease of 10 percentage points in the advantaged pass rate translates, on average, into a decrease in the advantaged/disadvantaged ELA gap of 4 percentage points. Put differently, when the advantaged pass rates *decline*, the gap *declines*. On the other hand, Graph V.4 indicates that *increases* in the disadvantaged fourth grade pass rate are associated with *reductions* in the gap. Here, 21% of the variation in the gap is accounted for by variation in the disadvantaged pass rate and an increase in the disadvantaged pass rate of 10 percentage points results in a reduction of the gap of an average of 4 percentage

points - schools with higher disadvantaged pass rates have lower gaps, but ones with higher advantaged pass rates have higher gaps.

Does school size matter? Graph V.5 indicates that there is no systematic relationship (R^2 = .01).[28] Two other school characteristics, percentage of students in poverty and percentage of students who are English language learners, are often thought to capture educationally challenging students and thus might be thought to be associated with larger gaps. Graphs V.6 and V.7 indicate, however, that the relationship between these attributes and the fourth grade gaps is *negative* (poverty with R^2 = .02) or insignificant (ELL with R^2 = .01). Apparently, these more difficult circumstances are not associated with higher gaps.

We next turn our attention to district-level characteristics that popular wisdom might suggest would be correlated with fourth grade ELA gaps. We analyze four of these: expenditures per pupil, the combined wealth ratio, the district percent of free and reduced lunch pupils, and district size as measured by enrollment. These relationships are displayed in Graphs V.8 through V.11. While expenditures per pupil, the combined wealth ratio and the percentage of free and reduced lunch students are not statistically significantly related to the gap, district enrollment is mildly negatively related (higher or larger is related to lower gaps; R^2 = .02). Math results in fourth grade are similar to ELA ones, as shown in Appendix Graphs V.1-10.

What do these relationships reveal? Several variables emerge that are related to the size of the ELA and math advantaged/disadvantaged gaps, including the subgroup pass rates, school poverty and ELL percentages, and district enrollment. Some of these relationships are weak and some are "perverse." For an example of the latter, higher advantaged pass rates are related to larger gaps as are larger percentages of students in poverty. Again, because there are lower gaps in New York City than in many other districts, these bivariate relationships reflect, disproportionately, the characteristics of the City.

2. Middle Schools. The relationships for middle schools are similar in direction to those of elementary schools, although the magnitude of these relationships is often somewhat stronger for the middle schools. Eighth grade results are shown in graphs V.12 through V.22, but we comment on only a few to illustrate the difference in strength between test score gaps and school and district characteristics relationships in the elementary and middle schools.

The relationship between the ELA and math test score gaps in middle schools are

[28] Note that as is true for all advantaged/disadvantaged analyses, New York City has much to do with the relationships. It will be important in future work to look within the City to see if these relationships hold, or put otherwise, to control for location with fixed effects.

stronger than in elementary (Graph V.12), as shown by the larger R^2 of .50 (compared to .32 for elementary schools) and correlation coefficient of .71 (compared to .57). Likewise the relationships with the overall and advantaged pass rates are stronger and larger in magnitude (see Graphs V.13 and V.14). The only relationship that is weaker here is between the disadvantaged pass rate and the gap (Graph V.15) - the R^2 is .10 (.21 in elementary schools).

Overall, more of the variables we examine are systematically related to the income gaps than to the race gaps, differing on occasion not only in significance but also in the direction of the relationships. For example, the overall pass rate is positively related to the income gap but negatively related to the race gap.

VI. School Test Score Gaps by Gender of Students

The last two columns of Table VI.1 (repeating results shown in Table III.1) show the male and female pass rates in the state on the ELA and on the math tests for fourth and eighth graders. The differences in these pass rates are the last ones we examine in this report. Note that females do better than males on the ELA in both grades, but in math, the small gaps in performance in fourth grade diverge somewhat in eighth grade, in favor of males. The organization of this part follows that of the previous section: Section A discusses the degree to which schools are mixed by sex, Section B analyzes the characteristics of mixed schools, Section C identifies schools with non-traditional, small, and traditional gaps along with their characteristics, and Section D uses regression analysis to study the relationship between gender gaps and school and district characteristics.

A. Gender Mix in New York State Schools

As before, we divide schools into three categories based upon their gender mix. Predominantly male (female) schools have more than five males (females) and five or fewer females (males). Mixed schools have more than five of each gender. New York State schools are basically mixed with respect to gender. As shown in Table VI.2, a trivial percentage of schools (between 0.9% and 2.1%) are single sex, and few students attend these (0.2%). Therefore, in this part, we study the gaps in mixed schools only. In addition, in contrast to our studies of race/ethnicity and income, the particularly interesting issues here seem to revolve around the change in gaps between fourth and eighth grades. Thus, for the most part, we present data on schools with both grades together on one table.

B. Test Score Gaps and Characteristics of Mixed Gender Schools

Table VI.3 shows test score gaps for both middle and elementary schools. These pass rates differ slightly from the ones reported in Table VI.1 for two reasons. First, the ones in Table VI.3 are school-level while the ones in VI.1 are student-level. That is, the ones in Table VI.3 use the school pass rates and average them across all schools, not accounting for differences in school size. The pass rates in Table VI.1, on the other hand, are based on all students in the state or, put differently, they are obtained by weighting the school pass rates with the number of students in a school of the appropriate gender. The second reason the rates differ in the tables is that Table VI.3 is based on mixed schools, excluding a very small number of single sex schools and their students.[29] Because this report analyzes *school* gaps, we use the unweighted numbers in analyses (weighted ones are included in Appendix tables).

On ELA assessments, in schools with fourth and with eighth grades, females outperform males, with a growing gap from 5.2 to 13.3 percentage points between the two grades. In math, on the other hand, the difference in elementary schools is trivial (0.4 percentage points), but grows in favor of males to 2.9 percentage points by eighth grade.

The characteristics of the mixed schools (Tables VI.4 and VI.5) are essentially equal to the characteristics of all schools with fourth and with eighth grades, since only the trivial number of single-sex schools are omitted from the analysis. The school location variables differ across the grades, probably reflecting differences in the size of schools, with New York City having relatively larger middle schools (and therefore a lower percentage of middle schools) and rural areas having relatively smaller sized middle schools (with a higher percentage of middle schools).[30] In terms of racial enrollment in the schools, middle schools show higher shares of whites and lower shares of non-whites, again perhaps due to larger schools in New York City and/or due to differences in dropout rates.

Table VI.5 shows school characteristics for mixed schools - eighth grade schools are larger, have lower poverty and English language learner rates, but are evenly divided male and female.

C. Examining the Gender Gap in Test Scores at the School Level

Table VI.6 shows the distribution of schools across categories defined by non-traditional, small or traditional test score gaps. Neither non-traditional (females above males) nor traditional

[29] This difference in rates holds for the race/ethnicity and for the income analyses as well, but the differences were less obvious there because of the larger number of "segregated" (and smaller number of mixed) schools.

[30] See Appendix Tables III.5 and III.6 for New York City and for all schools outside New York City.

(females below males) gaps seem desirable policy objectives, so we focus on the schools with small gaps, of which there are a sizeable number. For ELA tests, specifically, there are 35.3% and 17.1% of elementary and middle schools, respectively, with small gaps, while for math, the percentages are higher (44.8% and 45.3%).

The location of elementary schools by gap category, shown in Table VI.7, reveals that a disproportionate percentage of the schools with small gaps are located in the downstate suburbs and, in the case of math, in the upstate suburbs as well. For middle schools (Table VI.9), the distribution across locations differs, with New York City having a disproportionate share for both ELA and math, the Big 4 cities and upstate small cities joining New York City for ELA and the downstate suburbs joining New York City for math.

Do school characteristics differ by gap category? Data to answer this question for elementary schools are displayed in Table VI.8. Only a few characteristics stand out. Poverty is lower in schools with small gaps for both tests, and test scores, for both males and females on both tests, are for the most part higher than they are for schools with non-traditional or traditional gaps. In other words, the gaps are similar in small-gap schools (by definition), while scores of most groups are higher. Do we see the same patterns in middle schools? The statistics in Table VI.10 suggests the answer is "no." Middle schools that have small gaps educate more challenging students than other schools (higher percentages in poverty and who are English language learners) and they, by and large, have *lower* scores for males and for females than other schools. Thus in middle schools, in contrast to elementary schools, parity in test performance is achieved with a lowering of both group's scores. Recall that the middle schools with small gaps are disproportionately in New York City, while such elementary schools are disproportionately in the suburbs.

D. Statistical Relationships between Gender Gaps and District and School Characteristics

1. Elementary Schools. Graphs VI.1 through 11 display regression statistics and graphs of relationships between ELA male/female test score gaps and school and district variables for elementary schools, beginning with a graph of the relationship between the 4th grade male-female gap on the ELA and on the Math exam. To begin, the ELA and math gaps are positively correlated but not as tightly as the gaps in the race/ethnicity or income analyses. The relationship between the gap and the overall pass rate (Graph VI.2) is also positive but is so small as to be insignificant ($R^2 < .01$). The gender gap is positively correlated (R^2 of .10) with the male pass rate. As shown in graph VI.3, schools with higher ELA male pass rates on the ELA have larger

gaps. On the other hand, the relationship with the female pass rate is negative (Graph VI.4).

No systematic relationships are found between the gender gap and school-level variables-- neither enrollment (Graph VI.5), poverty (Graph VI.6) nor the percentage of English language learners are found to be significantly correlated with the gap (Graph VI.7).

Relationships with all four district characteristics -- expenditure per pupil (Graph VI.8), combined wealth ratio (Graph VI.9), percentage of free and reduced lunch participants (Graph VI.10), and total district enrollment (Graph VI.11) -- are of trivial magnitude and significance.

In math, the gender gap shows similar relationships as for ELA. For elementary schools, recall that on average the male/female gap in math is essentially zero.[31]

2. Middle Schools. Graphs VI.12 through VI.22 relate the male/female ELA test score gap and various school and district characteristics for middle schools. Once again, the ELA and the math gaps are related -- with a correlation of .42. Unlike elementary schools, middle schools show a slight negative relationship between the overall pass rate and the ELA gap (Graph VI.13, R^2 = .02), although the relationships with the male pass rate (positive, Graph VI.14) and the female pass rate (negative, Graph VI.15) are similar between elementary and middle schools.

School-level characteristics, including school enrollment (Graph VI.18), percentage of pupils in poverty (Graph VI.17), and percentage of pupils who are English language learners (Graph VI.18), are all positively related to the gender gap, although all with low correlations (R^2<.02). District characteristics are either unrelated to the gender gap (expenditures per pupil and the combined wealth ratio in Graphs VI.19 and 20, respectively) or very slightly positively related (percentage of students enrolled in free and reduced price lunch programs and total district enrollment in Graphs VI.21 and 22, respectively).

The relationships between the male/female gap in math and various characteristics sometimes differs from the ELA results and are more like the results for elementary schools. This is, perhaps, to be expected, because the direction of the male/female gap changes in math from negative (non-traditional) to positive (traditional), on average. In particular, the relationship with the overall pass rate is positive (r = .08, Appendix Graph VI.11), is highly positive with the male and mildly negative with the female pass rates, respectively, is mildly but insignificantly negative with all school characteristics (Graphs 12-16), and is negatively but mildly related to all district characteristics (Graphs 17-20).

[31] Relationships between male-female math test score gaps and school and district variables for the schools are shown in Appendix Graph VI.1-11.

VII. Summary and Conclusions

In this report, we have analyzed the gaps in test scores at the school-level between three groups of students in New York State schools with fourth and eighth grades – gaps in English and in math between racial/ethnic groups, between income groups, and between genders. While the aggregate gaps reflect familiar patterns, and many schools mirror these, there are some schools that manage to ameliorate the traditional gaps so that test performance between groups of students is comparable. After summarizing important conclusions, we then discuss next steps in this research.

Important Findings on Test Score Gaps in New York State Elementary and Middle Schools:

- There is a high degree of racial/ethnic segregation in New York State elementary and middle schools. About one-third of schools are predominantly white and these schools educate between one-fifth and one-quarter of the state's fourth and eighth grade students. Another fifth of the schools are predominantly non-white and educate between fifteen and twenty-three percent of the students.
- The gaps in test scores between racially segregated schools are over 2.5 times greater than the gaps in racially/ethnically mixed schools.
- The federal "No Child Left Behind" act requires that schools report gaps between racial subgroups if there are adequate numbers of students in each group. We chose a very low number of students tested (more than five) as our minimum cutoff, a number that will be quite a bit lower than New York or most other states choose.[32] Even with this low number, only 45.7% of New York State elementary schools (educating 51% of students) are in schools that would be required to report a gap. Well over half of the schools will not be required to confront gaps because they are effectively or "statistically" segregated, meaning there are too few children in the subgroup. Yet, the largest gap in performance is between segregated schools. How will the differences in overall pass rates be resolved with this system of accountability? Will reporting at a higher unit of aggregation (i.e., the district) take care of the problem? The answer in New York State is not likely because white schools are in predominantly white districts and non-white schools are in districts

[32] The lowest minimum subgroup sizes proposed for holding a school accountable are five in Maryland (and 95% confidence interval), 10 in Louisiana, South Dakota, and Utah (and a 99% confidence interval), 11 in New Hampshire (95% confidence interval), and 20 in New Jersey (and 95% confidence interval). New York State's current proposed minimum is 40 students (no confidence interval), although these figures are subject to change. See Lynn Olsen, "'Approved' Is Relative Term for Ed. Dept." *EdWeek*, August 6, 2003 <www.edweek.org/ew/ewstory.cfm?slug=43account.h22> and <www.edweek.org/ew/vol-22/43account.pdf> (August 6, 2003).

that have, on average, only 15.9% white students.

- Income segregation in New York State elementary and middle schools is not as high as racial/ethnic segregation, but is nonetheless significant. Well over one fifth of the schools and between fifteen and twenty-four percent of students attend schools that are economically homogenous, educating a student body that is either predominantly advantaged or disadvantaged.
- Again, the gap between income segregated schools is considerably greater than the gap within mixed income schools – well over two times greater.
- New York State has few single sex schools and only a trivial number of public school students attend a single sex school.
- The location and other characteristics of segregated schools differ from mixed schools.
  --Predominantly white schools are disproportionately in rural and upstate suburban districts, predominantly non-white schools are in New York City, and mixed schools are over-represented in downstate suburban districts. In addition, predominantly non-white schools educate a poorer and less English-proficient group of students.
  --Predominantly advantaged schools are disproportionately in upstate and downstate suburban districts, predominantly disadvantaged ones are in New York City and the other Big 4 New York State cities, and mixed schools are somewhat proportionately distributed compared to all schools. Not surprisingly, predominantly advantaged (disadvantaged) schools educate a less (more) poor and more (less) English-proficient student body.
- Between six and seven percent of schools exhibit comparable scores (small gaps) between race groups and between twelve and nineteen percent of schools exhibit parity in performance between income groups. (Tables IV.9 and V.9).
- These schools with small gaps for race are disproportionately located in New York City and downstate suburbs. (IV.10) In terms of income, they are disproportionately located, for elementary schools, in New York City, downstate small cities, and downstate suburbs (V.10) and for middle schools, in New York City, the other Big 4 cities and downstate small cities (V.12).
- In terms of race, the schools with small gaps are disproportionately Asian and not Black. (IV.11) In terms of income, it is the mixed schools that differ from the rest and they are smaller, with less challenged students in terms of poverty and English language, and they are more advantaged. (V.11 and V.13).
- Gender gaps favor females in ELA assessments and are small for math in fourth grade but slightly favor males in eighth grade. (VI.3).

- Between seventeen and forty-five percent of schools, depending on grade and test, exhibit small test score gaps by gender.
- The schools with small gender gaps are disproportionately located in downstate suburbs in fourth grade (VI.7), and in New York City and, for ELA, the other four big cities in eighth grade (VI.9).
- In elementary schools, the schools with small gender gaps educate disproportionately fewer poor students (VI.8); in middle schools, the schools with small gender gaps educate disproportionately more poor students (VI.10).
- Regression analyses reveal the following about the relationships between gaps in mixed schools and school and district characteristics.

  --Race/ethnicity gaps are related to the pass rates of each subgroup in the school. That is, increases in the non-white pass rate as a whole, and particularly increases in the black and Hispanic pass rates in particular, have moderately strong associations with reducing the overall white/non-white gap. No other variable analyzed is related.

  --Analyses of income gaps show that there are a few variables that are related to the size of the ELA and math advantaged/disadvantaged gaps, including the subgroup pass rates, school poverty and district enrollment. Some of these relationships are "perverse." For example, higher advantaged pass rates are related to larger gaps as are smaller percentages of students in poverty. Again, because there are lower gaps in New York City than in many other districts, these bivariate relationships may significantly reflect the characteristics of the City.

  --The gender gaps show significant relationships only with the overall pass rate (middle schools only) and male or female pass rates, and, like the race/ethnicity gaps, show no relationships with other school or district characteristics. In middle schools, the relationship between the overall pass rate and the gap is negative. For both grades, the male pass rate and the gap are positively related, while the female pass rate and the gap are negatively related.
- The three gaps, therefore, show some similar and some different relationships. Only income gaps are related to any school or district characteristics aside from subgroup pass rates. For the white/non-white gap, nonwhites do less well in segregated than in mixed schools, while white scores are similar in these two types of schools. For the advantaged/disadvantaged gap, however, advantaged students do better in segregated schools and disadvantaged students do worse in segregated schools compared to mixed schools. In other words, segregated schools do badly for non-white and disadvantaged

students, but only advantaged students do better in segregated schools. Segregation is correlated with poor performance for the lower scoring subgroup (although segregation may not "cause" this poor performance.[33])

Finally, while the data did not allow us to cross tabulate race/ethnicity, income and gender by student, we can tabulate the numbers and percentages of schools that are achieving small and other gaps on two and three dimensions. Tables VII.1 and VII.2 show these results. Predictably, although disappointing, there are few schools succeeding in all three areas, although some achieve parity across the board (16 and 25 in elementary schools and 1 and 8 in middle schools). If the criterion is even stricter, such as succeeding in both ELA and math, there is almost none[34]. If the successful ones must have over 10 black, Hispanic and white students each, then there are 4 and 1 in fourth grade and 0 and 5 in eighth grade (not shown here).

Further and Future Research

There are several unanswered questions raised in this report and several directions in which additional research is warranted. While it is a hopeful sign that many schools are able to achieve small gaps across races, income and gender, we do not know if these one-time occurrences are long-lasting, persistent characteristics of school. Will these same schools do well in years hence? This can only be answered by looking at subsequent years' data. We also do not know whether the suggestive results from the simple regression analyses will hold up in more sophisticated analyses – would the bivariate relationships between gap size and characteristics be stronger or weaker or even reversed if more variables were included and/or more flexible functional forms employed, for example? How important are the New York City schools in driving these relationships? Additional analyses to disentangle these separate effects would be interesting and important for policy making. More generally, it would be helpful to decompose the statewide gap in performance into that portion due to differences within districts (schools) and that portion due to differences between districts (schools).

---

[33] It may be that students who are low performing for reasons other than segregation go to school together, perhaps because families that are alike in their desire for educational achievement or in their ability to help their children achieve it, live in the same neighborhoods and attend their local schools.

[34] For middle schools, there are no schools that are non-traditional for both exams in all three categories (sex, race/ethnicity, and income). There is one school that exhibits small gaps in performance for both exams in all three categories and there are eight schools that have non-traditional or small gaps for both exams in all three categories.
For schools with fourth grades, there are no schools that are non-traditional for both exams in all three categories. There are two schools that have small gaps for both exams in all three categories and there are seventeen schools that have non-traditional or small gaps for both exams in all three categories.

More fundamentally, while these analyses document the distribution of the gaps, we have learned far too little about what causes the gaps and, while additional analyses of existing data would be revealing, some of the important factors may not (yet) be quantified. Quantitative analyses based upon more years of data, and perhaps, eventually, individual student data[35] that allowed identification of many student attributes (race, gender, income etc.) simultaneously, would aid testing of theoretical models and pursuit of causality through statistical methods. Case studies of successful schools would be a great way to begin the process of identifying elements of school curriculum, teaching philosophy, leadership, etc. that contribute to parity in performance between groups of students.

[35] New York State is pursuing a unique student identifier, planned for implementation by 2006.

References

Clotfelter, Charles T. (1999). "Are Whites Still 'Fleeing'? Racial Patterns and Enrollment Shifts in Urban Public Schools, 1987-1996" NBER Working Paper w7290 <http://papers.nber.org/papers/w7290> (July 18, 2003).

College Board, The (2001). *2001 College-Bound Seniors: A Profile of SAT Program Test Takers.* <www.collegeboard.com/sat/cbsenior/yr2001/pdf/NATL.pdf> (July 8, 2002).

Cook, Michael D. and William N. Evans (2000). "Families or Schools? Explaining the Convergence in White and Black Academic Performance." *Journal of Labor Economics* 18, no. 4 (October): 729-754.

Jacobson, Jonathan, Cara Olsen, Jennifer King Rice, Stephen Sweetland, and John Ralph (2001). "Educational Achievement and Black-White Inequality." National Center for Education Statistics, US Department of Education's Office of Educational Research and Improvement <http://nces.ed.gov/pubs2001/2001061.PDF> (July 17, 2003).

Jencks, Christopher and Meredith Phillips (1998). *The Black-White Test Score Gap*, Washington DC: Brookings Institution.

Mayer, Susan (2001). "How Did the Increase in Income Inequality between 1970 and 1990 Affect American Children's Educational Attainment?" *American Journal of Sociology* 107, no. 1 (July): 1-32.

Myers, Samuel L., Jr. (2001). "Racial Disparities in Minnesota Basic Standards Test Scores, 1996-2000." Hubert H. Humphrey Institute of Public Affairs, University of Minnesota <www.hhh.umn.edu/centers/wilkins/techrep.pdf> (July 18, 2003).

New York State Department of Education (2002). *New York State School Report Card.* <www.emsc.nysed.gov/repcrd2002/database.html> (July 8, 2002).

Orfield, Gary (2001). "Schools More Separate: Consequences of a Decade of Resegregation." Harvard Civil Rights Project <www.civilrightsproject.harvard.edu/research/deseg/separate_schools01.php> (July 17, 2003).

Shannon, G. Sue (2002). "Addressing the Achievement Gap: A Challenge for Washington State Educators." Washington State Office of Superintendent of Public Instruction <www.k12.wa.us/research/pubdocs/Achievementgapstudy.pdf> (July 18, 2003).