

Educational Policy

<http://epx.sagepub.com>

So Many Children Left Behind: Segregation and the Impact of Subgroup Reporting in No Child Left Behind on the Racial Test Score Gap

Leanna Stiefel, Amy Ellen Schwartz and Colin C. Chellman
Educational Policy 2007; 21; 527 originally published online Apr 24, 2007;
DOI: 10.1177/0895904806289207

The online version of this article can be found at:
<http://epx.sagepub.com/cgi/content/abstract/21/3/527>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:

Politics of Education Association

Additional services and information for *Educational Policy* can be found at:

Email Alerts: <http://epx.sagepub.com/cgi/alerts>

Subscriptions: <http://epx.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 14 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://epx.sagepub.com/cgi/content/refs/21/3/527>

So Many Children Left Behind

Segregation and the Impact of Subgroup Reporting in No Child Left Behind on the Racial Test Score Gap

Leanna Stiefel
Amy Ellen Schwartz
Colin C. Chellman
New York University

Although the No Child Left Behind Act was intended to help “all students meet high academic standards,” it is focused on subgroups of low-achieving students. The authors analyze the possible impact of the legislation’s requirement for performance reporting by racial subgroup in light of the considerable racial segregation in U.S. schools. In particular, using data on elementary and middle schools in New York State, the authors show that the schools are so highly segregated that more than half are too homogeneous to report test scores for any racial or ethnic subgroups. In addition, they show that the racial achievement gap is greatest across segregated schools rather than within integrated ones. The authors analyze the characteristics of schools that are and are not accountable for subgroups, finding that urban schools and large schools are particularly likely to be accountable, and conclude with implications for the reach of the law and for incentives for school segregation.

Keywords: *accountability; segregation; racial achievement gap*

Research studies and media headlines alike document the large test score gaps between students who are Black or Hispanic and those who are White or Asian (Jencks & Phillips, 1998). For example, recent results from the National Assessment of Educational Progress (NAEP) showed that nationwide, 87% of White students but only 54% of Black students performed at or above the “basic” level on the 2003 eighth grade reading exam. Similar

Authors’ Note: We thank the New York State Education Finance Research Consortium for funding previous work that led to this article. All analyses and conclusions reflect the views of the authors alone.

disparities exist across all grade levels between White and Asian students on one hand and Black and Hispanic students on the other (National Center for Education Statistics, 2003a, 2003b).

The No Child Left Behind Act (NCLB), signed by President Bush on January 8, 2002, has the declared intentions of helping "all students meet high academic standards" by 2014 and, more specifically, of increasing the academic performance of subgroups of students with traditionally low levels of achievement.¹ A key requirement of the legislation is that states administer annual tests in reading and math in Grades 3 through 8 and one high school grade and make available school-level performance data disaggregated by race and ethnicity, poverty, gender, special education, and English proficiency status. Furthermore, states must use sanctions to hold schools and districts accountable for their success in meeting adequate yearly progress (AYP) goals, set by the states, for both overall performance and performance in each subgroup.²

Notice, however, that the subgroup-reporting component of the law, which has received particular emphasis in the popular press as a means of reducing racial disparities, may not, in practice, provide the hoped-for results.³ The most important reason is that the largest test disparities in performance between race groups are between schools that are racially segregated rather than within schools that are racially integrated. And NCLB subgroup reporting brings no additional pressure to bear to address the significant gap in performance between racially segregated schools.

In addition, the increased attention to the performance of minority students that emerges from the subgroup accountability will be unevenly distributed across schools, falling disproportionately on urban schools and substantially bypassing rural and suburban schools, which are more likely to be racially homogeneous.⁴ Finally, the implementation of minimum subgroup sizes for subgroup reporting (to ensure statistically appropriate subgroup performance scores) will create an incentive for schools, and perhaps districts, to manipulate the compositions of their student bodies to avoid the requirement to report performance in some subgroups. The implication is that the desire to avoid any additional pressure of NCLB subgroup accountability, beyond the pressure created by the overall AYP standards, may mean that school segregation may be exacerbated and the race gap worsened.⁵

We used data on elementary and middle schools in New York State to illustrate these phenomena, focusing on subgroups of White and non-White students. To preview the results, we found that school segregation in New York State significantly limits the impact that subgroup reporting can possibly have on lowering the racial test score gap. Put simply, a surprisingly

small share of all New York State schools will be held accountable for disparities in performance between Whites and non-Whites. Moreover, the gap in performance between Whites and non-Whites in racially segregated schools is considerably larger than within integrated schools. The new subgroup reporting provisions of NCLB will apply where the gap is less severe. Finally, New York State schools that will be held accountable for racial subgroups are distinguished from ones that will not by their larger sizes and their concentration in urban areas.

We consider whether raising accountability for subgroup performance to the district level mitigates the effects of segregation on NCLB subgroup accountability measures. Unfortunately, segregation across districts is also quite profound, and district accountability for the race gap is found to result in far less than the full coverage of schools. Our results raise important questions about whether the federal NCLB subgroup accountability will act in such a way as to effect higher test scores and reduce disparities in performance between race groups, or whether school segregation in the United States will render school and district subgroup accountability measures ineffective.

Review of Related Literature

The Race Gap in Test Scores

There is considerable evidence of the persistence of gaps in performance between White and non-White students (and specifically between White and Hispanic and between White and Black students) using both state and national data on standardized test scores from early childhood to high school. Jencks and Philips's (1998) volume *The Black-White Test Score Gap* presents evidence and several studies that examined the scope and causes of the disparity. Hedges and Nowell (1998), for example, used high school test score data from six nationally representative surveys conducted between 1965 and 1992 to find that the test score gaps between Whites and Blacks declined over this period, that adjusting for socioeconomic factors reduced gaps by less than half, and that the trends in adjusted and unadjusted test score gaps followed similar paths. More recently, Fryer and Levitt (2004) found that adjusting for the background characteristics of students reduced the Black-White achievement gap to zero at kindergarten entry (but not thereafter), on the basis of the cohort of kindergarteners through first graders from the Early Childhood Longitudinal Study.

Several studies have made efforts to identify the extent to which disparities in test scores reflect differences within schools rather than between

schools. Cook and Evans (2000), for example, decomposed reductions in the White-Black NAEP test score gap among middle school students between 1970 and 1988. They found that 25% of the change could be attributed to family and school characteristics, whereas 75% could be attributed to within-school reductions. Reardon (2003), using the same Early Childhood Longitudinal Study data set as Fryer and Levitt (2004), found that during kindergarten, test score gaps grow within schools, even after controlling for socioeconomic status and school context, and in first grade, between-school rather than within-school components explain growth in gaps. Stiefel, Schwartz, and Ellen (2004) decomposed New York City fifth and eighth grade reading test score gaps (White-Black, White-Hispanic, and White-Asian in school years 1995-1996 and 2000-2001) into differences attributable to students' socioeducational traits, within-school factors, and between-school factors. They found that for both Black and Hispanic students, the between-school effects accounted for approximately one third of the total difference in scores.

Segregation

School segregation has a long history in the United States, linked in part to residential segregation. Massey and Denton (1993) provided a comprehensive examination of residential segregation during the past century, describing how racial divisions between residential communities were maintained even during an era of great population shifts, with Black population movement to northern cities and White migration to suburbs across the nation. They illustrated that residential segregation became a permanent feature of U.S. cities after World War II and into the 1980s: The dissimilarity index remained in the high 80s from 1940 to 1970 in the 30 metropolitan areas with the largest Black populations, and the isolation index did not dip below 50 in any city from 1970 to 1980.⁶ More recently, Orfield and Lee (2004) examined school segregation, exploring student segregation across schools within the nation and within and between states. They found a trend toward resegregation between 1991 and 2001 in states that had been highly desegregated in 1991. Interestingly, in 2001, New York was one of the most segregated states for Black and Hispanic students according to two different measures: exposure to White students and percentage minority in majority White schools. Furthermore, Clotfelter (2001) found that school racial segregation is increasing, although within-district segregation is declining, leaving most segregation attributable to between-district sorting.

Accountability and Testing

Despite the absence of consensus on the causes of test score gaps or the best methods to reduce those disparities, NCLB codified the belief that school and district accountability would reduce achievement gaps within each state and across the nation.

The notion that accountability measures can reduce test score gaps (or improve students' performance) has considerable appeal and some supporting evidence (Hanushek, 2002; Loeb & Strunk, 2003).⁷ Ladd (1999) suggested that the accountability system implemented by the Dallas Independent School District led to improvement in performance for Hispanic but not Black students on state tests relative to students in other large Texas cities. In contrast, Smith and Mickelson (2000) found that an accountability program in Charlotte-Mecklenburg, North Carolina, did not increase students' performance relative to the state average. Carnoy and Loeb (2003) used NAEP scores and a ranking of states' accountability systems to assess the effects of systems that vary across, not within, states. They found that states with stronger accountability systems had greater gains in fourth and eighth grade scores for Hispanic and Black students and across-the-board eighth grade gains for students of all races. Fuller and Johnson (2001) found similar results using data from Texas on state standardized exams. Finally, Hanushek and Raymond (2005), using NAEP scores, concluded that accountability systems improve student performance overall while narrowing Hispanic-White but not Black-White achievement gaps.

In a different vein, Kane and Staiger (2003) have explored some of the difficulties anticipated in the implementation of accountability measures, most recently in Peterson and West's (2003) *No Child Left Behind?*⁸ They found that accountability schemes requiring schools to ensure test score gains by students from all racial and ethnic backgrounds tend to penalize integrated schools by making year-to-year measures of achievement changes more volatile. These subgroup accountability provisions fall more heavily on integrated schools even when, on average, over several years, their students are performing better than students in segregated schools.

Stiefel, Schwartz, and Chellman (2003) pursued a somewhat different direction, analyzing the potential effects of subgroup size exemptions on the extent to which accountability for both White and non-White student performance will bind on schools in New York State. Here, we build on this research and their suggestion that segregation will limit the effectiveness of the subgroup accountability parts of reforms in reducing test score gaps because of the presence of racially homogeneous schools.

Data

We used school- and district-level data obtained from the New York State Education Department (NYSED) for 2001-2002.⁹ These data included school-level scores for tests in English language arts (ELA) and math and school characteristics, such as enrollment and percentage of students eligible for free and reduced-price lunch. We combined these data with district-level variables obtained from the NYSED, including expenditures per pupil, racial breakdown at the district level, and percentage of students in the district eligible for free and reduced-price lunch.

A small number of schools were eliminated from the analyses because they educate large proportions of special education students. Specifically, these schools have more than 40% of students with individualized education programs, and most (75%) are located in New York City's District 75, a program-specific (rather than geographically determined) district serving mostly full-time special education students.¹⁰

Schools educating fourth graders are largely elementary schools, schools educating eighth graders are largely middle schools, and they are referred to as elementary and middle schools, respectively, below. Note, however, that some schools serving eighth grade students also contain high school grades, and some schools (211) serve both fourth and eighth graders. These latter schools appear in both the elementary and middle school analyses.

Table 1 shows descriptive statistics for elementary and middle schools in 2001-2002. There were 2,249 elementary schools and 1,068 middle schools in New York State with 10 or more students tested in ELA and math in school year 2001-2002. This cutoff and the exclusion of schools with high proportions of special education students increased the reliability of the subsequent school gap analyses, although results using all schools were qualitatively the same.

The smallest elementary schools in the state tested, by construction, 10 students in fourth grade, whereas the largest tested 370, with a statewide average of 93 students tested in fourth grade. The average middle school tested 188 students in eighth grade, whereas the largest tested 969 (the minimum, again by definition, was 10). These differences in grade sizes also were reflected in total school enrollment: The state average elementary school enrollment was 577 (ranging from 58 to 2,164), whereas the average middle school enrolled 741 students (ranging from 102 to 2,532).¹¹

In 2001-2002, approximately 12% of students attending school in New York State were in special education programs, about 5% were English-language learners, and approximately 44% were poor. The NYSED divides the state's districts into geographic groups for reporting purposes. Using

Table 1
Descriptive Statistics for New York State Elementary and Middle Schools, 2001-2002

Variable	Elementary Schools (n = 2,249)			Middle Schools (n = 1,068)		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
Grade level						
Total tested, ELA	93.2	10.0	370.0	188.2	10.0	969.0
Percentage of fourth and eighth grade students tested						
White	60.4	0.0	100.0	64.3	0.0	100.0
Non-White	39.6	0.0	100.0	35.7	0.0	100.0
School level						
School enrollment	577	58	2,164	741	102	2,532
Percentage of school enrollment						
English-language learner	5.7	0.0	49.8	4.3	0.0	78.4
Special education	11.0	0.0	38.9	13.3	0.0	39.3
In poverty	45.0	0.0	100.0	42.4	0.0	100.0
District level						
Location						
New York City	29.6			27.4		
Big 4 cities (not New York City)	6.0			6.3		
Rural	10.2			16.4		
Downstate small cities	1.4			0.8		
Upstate small cities	10.1			6.6		
Downstate suburbs	20.1			16.1		
Upstate suburbs	22.6			26.4		

(continued)

Table 1 (continued)

Variable	Elementary Schools (<i>n</i> = 2,249)			Middle Schools (<i>n</i> = 1,068)		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
District size (total enrollment)	317,605	108	1,057,682	294,718	102	1,057,682
Expenditures per pupil	\$12,329	\$8,421	\$27,916	\$12,382	\$8,421	\$25,560
Property wealth per pupil	\$260,100	\$48,223	\$9,578,847	\$256,476	\$48,223	\$2,957,663
Income per pupil	\$101,409	\$22,746	\$953,732	\$94,434	\$22,746	\$682,495

Note: The sample included all students from schools with 10 or more students tested in English language arts (ELA) and math and excluded schools in which more than 40% of the students had individualized education programs (or special education). Descriptive statistics are not student weighted. The location variable includes regions of New York State as defined by the New York State Education Department; every district is categorized by the state into only one region.

these groups, New York City has the largest share of schools in the state, approximately 30% of all elementary schools and 27% of all middle schools, followed by the upstate suburbs, with nearly 23% of elementary schools and 27% of middle schools. The “big 4 cities” (Buffalo, Rochester, Syracuse, and Yonkers) represent 6% of elementary and middle schools overall statewide. The smallest fraction of schools is located in downstate small cities, with 1.4% of elementary and 1.0% of middle schools.

Elementary and middle schools are in districts with similar average district expenditures per pupil (about \$12,500) and similar property wealth per pupil (about \$258,000). Income per pupil is slightly different, comparing across school levels, with elementary schools in districts with \$101,409 and middle schools in districts with \$94,434.

Empirics

How Large Are Disparities in Statewide Test Scores by Race?

Table 2 shows that there are large disparities in test score results between White and non-White students¹² in New York State. In 2001-2002, White and Asian pass rates were above the state average on both exams and grades, whereas Black, Hispanic, and American Indian scores were below the average. White students outperformed non-White students (including Asians) by 27.7 percentage points on the fourth grade ELA exam and by 29.1 points on the fourth grade math exam. Variation by specific racial subgroup was large: Asian students did as well as White students on the ELA exam while outperforming White students on the math exam; Black students were the lowest performing, with a 42.4% pass rate on ELA and 45.8% pass rate on math.

The results for eighth grade were similar to those for fourth grade: White students outperformed non-White students by 30.7 points on the ELA exam and 34.8 points on the math exam. Again, variation by specific racial subgroup was large: Asians outperformed White students on both ELA and math, whereas Black students continued to perform the worst, with an average pass rate of 21.5% on the eighth grade ELA exam and 21% pass rate on the math exam.¹³

Which Schools Are Held Accountable for Test Score Gaps by Race?

Reflecting the understanding that performance measures are volatile and unreliable when based on small sample sizes, NCLB legislation made clear

Table 2
Student Pass Rates, New York State English Language Arts (ELA) and Math Tests,
Fourth and Eighth Grades, 2001-2002, Total and by Subgroup

Variable	Total	White	Non-White	Black	Hispanic	Asian	American Indian
Fourth grade							
ELA pass rate	61.9	74.2	46.5	42.4	42.7	74.1	45.1
Number of students	209,699	116,810	92,889	43,150	37,303	11,551	885
Number of schools	2,249	2,073	2,009	1,716	1,693	1,368	445
Math pass rate	68.0	81.1	52.0	45.8	49.6	83.2	56.1
Number of students	213,388	117,194	96,194	43,117	40,467	11,722	888
Number of schools	2,249	2,069	2,012	1,717	1,706	1,386	444
Eighth grade							
ELA pass rate	44.8	57.5	26.8	21.5	22.5	57.8	30.7
Number of students	200,945	117,942	83,003	38,685	32,623	10,948	747
Number of schools	1,068	1,007	969	834	822	695	292
Math pass rate	48.2	62.9	28.1	21.0	22.9	67.1	36.8
Number of students	203,918	117,797	86,121	38,777	35,108	11,480	756
Number of schools	1,068	1,004	969	833	827	696	305

Source: Test score data are from New York State School Report Card (New York State Education Department, 2002).

Note: The sample included all students from schools with 10 or more students tested in ELA and math and excluded schools in which more than 40% of the students had individualized education programs.

Table 3
Distribution of New York State Schools and Students Tested
by Racial Accountability Status of School, 2001-2002

Variable	Schools		Students	
	<i>n</i>	%	<i>n</i>	%
Fourth grade				
Schools, White-accountable only	953	42.4	72,484	34.6
Schools, non-White-accountable only	523	23.3	55,503	26.5
Schools accountable for both groups	773	34.4	81,712	39.0
Total	2,249	100.0	209,699	100.0
Eighth grade				
Schools, White-accountable only	392	36.7	42,027	20.9
Schools, non-White-accountable only	193	18.1	35,509	17.7
Schools accountable for both groups	483	45.2	123,409	61.4
Total	1,068	100.0	200,945	100.0

Source: Test score data are from New York State School Report Card (New York State Education Department, 2002).

Note: The sample included all students from schools with 10 or more students tested in English language arts and math and excluded schools in which more than 40% of the students had individualized education programs. White-accountable schools had test performance data for 10 or more white students tested and fewer than 10 non-White students tested. Non-White-accountable schools had test performance data for fewer than 10 White students tested and 10 or more non-White students tested. The number of students is based on the availability of performance data for English language arts. Numbers would be slightly different if based on data for math.

that schools with small numbers of students in subgroups would not be held accountable for those subgroups' performance. No specific threshold number of students was specified, however.¹⁴ We set a threshold requiring 10 or more test scores in a subgroup (White or non-White) because current plans show New York State using a threshold of 30 students over three grades, or an average of 10 per grade.¹⁵ Table 3 shows the distribution of schools and students tested by racial accountability "status": whether a school will be held accountable for the performance of both White and non-White students. For elementary schools, 42.4% (representing 34.6% of fourth grade students) would be held accountable for White students only, whereas 23.3% of schools (representing 26.5% of fourth graders) would be accountable for non-White students only. That leaves 34.4% of schools (with 39% of students) that would be held accountable for both White and non-White scores and that could report subgroup scores.

Similar trends were found for middle schools, among which 36.7% (representing 20.9% of eighth grade students) would be held accountable for White students only, whereas 18.1% (educating 17.7% of non-White students) would be held accountable for non-White students only. The remaining 45.2% of schools (educating 61.4% of eighth grade students) would be held accountable for both White and non-White students. Again, schools accountable for both groups are larger.

Fewer than half of the schools in the state would be accountable for the performance of both White and non-White students; this pattern reflects the legacy of residential segregation in New York State (Orfield & Lee, 2004). We found considerable segregation across the schools in the state. Statewide, in elementary schools, the dissimilarity index was 0.76 for school year 2001-2002, indicating that 76% of non-White students would have to transfer from schools in which they are overrepresented to schools in which they are underrepresented to achieve a distribution in each school proportional to their statewide percentage representation.¹⁶ In middle schools, the dissimilarity index was 0.74 in 2001-2002. For schools accountable for both White and non-White students, the index fell to 0.49 in elementary schools and 0.59 in middle schools. Although these values were smaller than for elementary and middle schools as a whole, even those schools that were accountable for both groups were still highly segregated.

What Are the Correlates of School Subgroup Accountability Status?

Table 4 shows the ELA and math exam results for elementary schools by the accountability status of schools. The first and second columns include those schools that were accountable only for White students or only for non-White students, respectively. The third column combines schools from the first two columns into an aggregated single-group (White-only plus non-White-only) accountability column, and the fourth column includes schools that were accountable for both White and non-White students.

ELA pass rates for White-only accountable and non-White only accountable schools were lower across all groups compared with pass rates in schools that were accountable for both groups. The total pass rate on the fourth grade ELA exam was 59.1% in single-group-accountable schools compared with 66.3% in schools accountable for both groups. The pass rate for White students on the fourth grade ELA exam was very similar across groups: a 73.5% pass rate in single-group-accountable schools compared with 75.2% in schools accountable for both groups. The non-White pass

Table 4
Fourth Grade Student Pass Rates and Gaps by Racial Accountability
Status of New York State Schools, 2001-2002

Variable	White-Only Accountable (953)	Non-White-Only Accountable (523)	White-Only or Non-White-Only Accountable (1,476)	Accountable for Both Groups (773)
English language arts				
Total pass rate	73.5	40.2	59.1	66.3
Number of schools	953	523	1,476	773
White pass rate	73.8	—	73.5	75.2
Number of schools	953	—	1,300	773
Non-White pass rate	—	39.9	41.3	54.8
Number of schools	—	523	1,236	773
Gap	—	—	32.2	20.4
Number of schools	—	—	—	773
Math				
Total pass rate	80.9	45.5	65.3	72.1
Number of schools	953	523	1,476	773
White pass rate	81.2	—	80.9	81.3
Number of schools	953	—	1,296	773
Non-White pass rate	—	45.2	46.6	60.7
Number of schools	—	523	1,239	773
Gap	—	—	34.3	20.6
Number of schools	—	—	—	773

Source: Test score data are from New York State School Report Card (New York State Education Department, 2002).

Note: The sample included all students from schools with 10 or more students tested in English language arts and math and excluded schools in which more than 40% of the students had individualized education programs.

rate, however, was 41.3% in single-group-accountable schools compared with 54.8% in schools accountable for both groups. Non-White students performed better in schools that were accountable for both groups, compared with schools that were accountable for only White students¹⁷ or accountable for only non-White students, whereas White students performed about the same in all groups. In other words, schools that were even somewhat “integrated” contained students (White or non-White) who performed better than schools that were segregated. This result is amplified when considering test score gaps; in schools accountable for both groups, there were smaller test score gaps between White and non-White students. There was a 32.2-point gap across single-group-accountable schools compared with a

Table 5
Eighth Grade Student Pass Rates and Gaps by Racial Accountability
Status of Schools, 2001-2002

Variable	White-Only Accountable (392)	Non-White-Only Accountable (193)	White-Only or Non-White-Only Accountable (585)	Accountable for Both Groups (483)
English language arts				
Total pass rate	53.4	19.7	38.0	49.1
Number of schools	392	193	585	483
White pass rate	53.6	—	53.4	59.8
Number of schools	392	—	524	483
Non-White pass rate	—	19.6	20.3	31.7
Number of schools	—	193	486	483
Gap	—	—	33.0	28.1
Number of schools	—	—	—	483
Math				
Total pass rate	60.4	20.4	41.7	52.2
Number of schools	392	193	585	483
White pass rate	60.7	—	60.4	64.2
Number of schools	392	—	521	483
Non-White pass rate	—	20.3	21.1	33.5
Number of schools	—	193	486	483
Gap	—	—	39.3	30.7
Number of schools	—	—	—	483

Source: Test score data are from New York State School Report Card (New York State Education Department, 2002).

Note: The sample included all students from schools with 10 or more students tested in English language arts and math and excluded schools in which more than 40% of the students had individualized education programs.

20.4-point gap in schools accountable for both groups. We found similar results for fourth grade math and for both tests in middle schools (Table 5).¹⁸

These findings indicate that more of the statewide test score gap is due to between-school differentials than to within-school differentials. To quantify the between- and within-school components, we decomposed the statewide test score gap into between and within components as follows:

$$\begin{aligned} \bar{X}^W - \bar{X}^N &= \left[\frac{\sum N_{ws} \bar{X}_s^w}{\sum N_{ws}} - \frac{\sum N_{ns} \bar{X}_s^w}{\sum N_{ns}} \right] \\ &= \left[\frac{\sum_s N_{ws} (\bar{X}_s^w - \bar{X})}{\sum_s N_{ws}} - \frac{\sum_s N_{ns} (\bar{X}_s^w - \bar{X})}{\sum_s N_{ns}} \right] + \left[\frac{\sum N_{ws} \bar{X}_s}{\sum N_{ws}} - \frac{\sum N_{ns} \bar{X}_s}{\sum N_{ns}} \right], \quad (1) \end{aligned}$$

Table 6
Decomposition of Test Score Gaps, 2001-2002

	Fourth Grade				Eighth Grade			
	ELA		Math		ELA		Math	
Statewide gap	27.7	100%	29.0	100%	30.8	100%	34.8	100%
White deviation from mean school performance	1.9		2.0		2.7		2.8	
Non-White deviation from mean school performance	-2.3		-2.5		-3.9		-3.8	
Average within-school gap	4.2	15%	4.5	16%	6.7	22%	6.6	19%
White-weighted school pass rate	72.3		79.0		54.8		60.0	
Non-White-weighted school pass rate	48.8		54.5		30.7		31.9	
Average between-school difference	23.5	85%	24.5	84%	24.1	78%	28.1	81%

Note: The sample included all students from schools with 10 or more students tested in English language arts (ELA) and math and excluded schools in which more than 40% of the students had individualized education programs. White deviation from mean school performance was calculated as the difference between the White-weighted White pass rate and the White-weighted school pass rate. Non-White deviation from mean school performance was calculated as the difference between the non-White-weighted non-White pass rate and the non-White-weighted school pass rate. Average within-school gap was calculated as the White deviation from mean school performance minus the non-White deviation from mean school performance. Average between-school gap was calculated as the non-White-weighted school pass rate subtracted from the White-weighted school pass rate. These weighted pass rates can be interpreted as the state average pass rate if either non-White students (who are disproportionately educated in poorly performing schools) or White students were evenly distributed across the state and those redistributed students took on their new schools' averages.

where the statewide test score gap is calculated as \bar{X}^W , the statewide White student pass rate (the White-weighted White pass rate in the state), minus \bar{X}^N , the statewide non-White student pass rate (the non-White-weighted non-White pass rate in the state), and N_{ws} is the number of White students in a school and N_{ns} is the number of non-White students in a school. The within-school portion of the equation includes $\bar{X}_s^w - \bar{X}_s$ as the White deviation from the school mean and $\bar{X}_s^n - \bar{X}_s$ as the non-White deviation from the school mean.

Table 6 shows that 85% of the total statewide gap in fourth grade ELA results was attributable to between-school differences. Similar results were

calculated for fourth grade math and eighth grade ELA and math exams, also shown in Table 6. Thus, a much larger disparity exists between segregated schools than within mixed schools, yet only the latter schools will be held accountable for both White and non-White student performance.

What Are the Characteristics of Schools Accountable for Both Race Groups?

We used probit regression analysis to understand which characteristics of schools and districts were associated with having adequate numbers of both White and non-White students to report subgroup scores. These regressions are available on request; here, we report results. All else equal, schools that were larger; located in larger districts; located in districts with higher percentages of Hispanic and Asian or American Indian students; and in downstate suburbs, large cities, or New York City were more likely to be accountable for subgroups. Thus, the subgroup accountability will be unevenly distributed, falling on already fiscally stressed cities.

Are Districts Accountable When Schools Are Not?

An important provision of NCLB ensures that if a school has too few students in a subgroup to be held accountable for that group, the district will be held accountable, allowing for small subgroup numbers in individual schools to be aggregated to the district level. How does this provision affect accountability for small racial subgroups at the school level? When we aggregated subgroup numbers to the district level, we found that in 2001-2002, 261 of 657 districts (39.7%) were accountable for both White and non-White fourth grade students. Table 7 shows that district-level accountability will capture 920 more elementary schools and 84,860 more students (30,040 more White students and 54,820 more non-White students) than school accountability alone. That leaves 556 schools (24.7%) that will remain unaccountable even when districts are held accountable (396 districts, or 60.3%, will not be held accountable for both White and non-White students).

At the middle school level, 257 of 635 districts (40.5%) were accountable for eighth grade White and non-White students. District-level accountability will capture 202 more middle schools and 36,868 more students (2,815 White and 34,053 non-White) than when schools alone are held accountable. That leaves 382 schools (35.8%) that will remain exempt from accountability even at the district level (378 districts, or 59.4%, will not be held accountable for both White and non-White students).¹⁹

Table 7
Additional Schools (Students) Covered by
Accountability at District Level, 2001-2002

Variable	Schools		Students	
	<i>n</i>	%	<i>n</i>	%
Fourth grade, accountability coverage				
School level for White and non-White	773	34.4	81,712	39.0
District level for White and non-White	920	40.9	84,860	40.5
Exempt from accountability for White and non-White	556	24.7	43,127	20.6
Total	2,249	100.0	209,699	100.0
Eighth grade, accountability coverage				
School level for White and non-White	483	45.2	123,409	61.4
District level for White and non-White	202	18.9	36,868	18.4
Exempt from accountability for White and non-White	382	35.8	40,668	20.2
Total	1,068	100.0	200,945	100.0

Source: Test score data are from New York State School Report Card (New York State Education Department, 2002).

Note: The sample included all students from schools with 10 or more students tested in English language arts and math and excluded schools in which more than 40% of the students had individualized education programs. White-accountable schools had test performance data for 10 or more White students tested and fewer than 10 non-White students tested. Non-White-accountable schools had test performance data for fewer than 10 White students tested and 10 or more non-White students tested. The number of students is based on the availability of performance data for English language arts. Numbers would be slightly different if based on data for math.

What kinds of districts are accountable for both White and non-White students? We found results very similar to our analysis at the school level.²⁰ Large total district enrollment increases the likelihood of accountability. Rural and upstate suburban districts are less likely to be accountable, whereas districts with higher percentages of Hispanic and Asian or American Indian students are more likely to be held accountable. We found similar patterns for middle schools and across years.

Although future work will better assess the actual behavioral response to the unintended incentives set up by NCLB, to get a sense of the possible

magnitude, we calculated the number of schools that most easily could manipulate student enrollment or test taking to avoid the 10-student minimum. These schools were defined as having greater than or equal to 7 but less than or equal to 13 White or non-White students. We found that of the 2,249 elementary schools in the state (the total number of schools from Table 1 that had tested 10 or more students in both ELA and math), 374 (16.6%) were near the cutoff of 10 White or non-White students. In addition, of the 374, only 44.4% were considered to be accountable for both groups in our previous analysis; in other words, those are the schools in which we are most likely to see the behavioral response to get under the cutoff of 10 students to avoid accountability measures.

In addition, of the 1,068 middle schools in the state, 119 (11.1%) were near the cutoff of 10 students. Of those 119, 48% were considered to be accountable in our previous analysis, and thus have the greatest incentive to get under the 10-student minimum to avoid accountability. These schools with the greatest incentive to get under the accountability minimums are, on average, smaller, less poor, much less likely to be in New York City, and much more likely to be in downstate suburbs than the schools in the rest of New York State.

We emphasize that these estimates are conservative, because we used a low cutoff of 10 students for racial subgroups and we defined racial subgroups broadly (White and non-White), as opposed to the five specific race groups included in the data (White, Black, Hispanic, Asian, and American Indian).

Conclusions

Evidence in this article points to a number of limitations in the NCLB subgroup accountability system and its ability to lower racial test score gaps in the face of segregation in New York State. First, it is clear that setting a threshold subgroup size high enough for statistically meaningful performance measures is likely to require excluding a large number of schools—and, it is important to note, districts—from accountability for their subgroups' performance levels. If the intention is to compare performance between race groups, even using highly aggregated groupings such as White and non-White students, the coverage is far from complete. Incomplete coverage reduces the reach of sanctions in the NCLB subgroup reporting reforms, distributes subgroup accountability measures unequally across schools and districts, and possibly creates incentives for schools and districts to

remain below accountability thresholds by manipulating enrollment or test taking.

Second, the socioeconomic characteristics of schools held accountable for subgroups differ from those that will be exempted from such accountability. Large, urban schools that are more racially and ethnically integrated will bear a disproportionate accountability burden, whereas rural schools, and to a lesser degree upstate suburban schools, that are less integrated will be disproportionately exempted. Put differently, racial subgroup accountability may in practice affect an atypical set of schools: disproportionately urban, Hispanic or Asian, and large rather than rural or suburban, White, and small. To a large extent, these results reflect the continuing legacy of residential and school racial segregation in New York State.

Perhaps more disturbing is the possibility that the desire to avoid the pressure of subgroup accountability may serve to discourage school integration, if districts (or schools) are able to manipulate enrollment or test-taking behavior to drop below subgroup reporting thresholds designed to ensure statistically valid performance measures.

Third, the schools accountable for subgroup reporting are not the ones with the lowest test scores for non-White students: Average scores are consistently lower among the exempted schools. The NCLB subgroup accountability provision will bind most in schools and districts that are contributing less to statewide test gaps. The role of subgroup reporting in NCLB cannot be expected to eliminate racial test score disparities, because the largest disparities lie between schools without significant subgroup populations. Thus, to the extent that subgroup accountability is intended to raise test scores, the results suggest that accountability measures may be missing an important segment of the population and will, then, have a lesser impact than had been imagined.

Fourth, more complete subgroup accountability coverage can be obtained by raising accountability to the district level. Ultimately, it may be necessary to raise subgroup accountability sanctions to the state level to shine a light on the disparity in performance between subgroups of all students. Without the full weight of subgroup accountability falling across all districts and schools, the provisions of NCLB cannot be expected to achieve its goals even with more stringent AYP goals in later years.

In the end, closing racial test score gaps will have to rely more on the efficacy of the AYP provisions of NCLB than on the subgroup reporting provisions. If AYP provisions are maintained to include 100% of students and if standards of proficiency are not lowered, then as 2014 approaches, all schools, segregated or not, will be "required" to bring all students to

proficiency. As a number of analysts have shown, however, the likelihood of meeting this goal is quite low (Linn, 2005), and states may well be tempted to lower proficiency standards. Without strict adherence to AYP, however, racial gaps will not be closed through subgroup reporting.

Notes

1. In a letter to the editor of the *New York Times*, Rod Paige (2003), President Bush's first secretary of education, wrote, "In the past, a lot of lip service has been given to closing the achievement gap—the law's central tenet—but now the framework is set for real reform to finally happen" (p. A20).

2. More precisely, schools receiving federal aid that fail to achieve AYP in their students' performance are, depending on the number of years of failure, listed as "in need of improvement," required to provide tutoring for students with inadequate performance, required to let students transfer to better performing schools, or required to "reconstitute" (U.S. Department of Education, 2004b).

3. An article on the Princeton, New Jersey, school district illustrates the emphasis on the subgroup reporting component of the law by the popular press (Freedman, 2005). The article quoted the co-chair of the school district's minority education committee: "I do have some suspicions when measurements come from standardized tests alone. But if it's going to shine a bright light on the inadequacies of the system, especially as it regards children of color, then I'm all in favor" (p. B7). Over time, the AYP requirements of NCLB would impose 100% proficiency for all students and therefore apply to all schools, including racially segregated ones. However, scholars have discussed how difficult it will be to achieve 100% proficiency by 2014 and how NCLB requirements may be changed by then (e.g., Linn, 2005).

4. Although Black students constitute 17.2% of public school students enrolled in the United States, they make up 32.0% of public school enrollment in the large cities of the United States. Similar disparities exist for Hispanic students (16.3% in the United States compared with 32.0% in large cities) and Asian and Pacific Islander students (4.1% compared with 28%).

5. The notion that apparently neutral federal policies can exacerbate segregation is not new. For example, Voith and Gyourko (2002) examined the extent to which the tax treatment of owner-occupied housing leads to the decentralization of land use, such that higher income households become more separated from lower income individuals, who become isolated in central cities. Similarly, Persky and Kurban (2003) found that federal subsidies reduce the price of land outside cities, providing an incentive to those who can take advantage of the subsidies to leave cities.

6. The dissimilarity index captures the overall unevenness of a population's distribution and is used extensively in residential and school segregation literatures; this index describes the percentage of one group of individuals who would have to move to be represented in all areas equal to their population proportion. The isolation index is a type of interaction index that shows how likely an individual is to interact with individuals from his or her own group.

7. We review the literature on the effects of improving student test scores, through sanctioned methods, in response to accountability systems. For further background on school strategies that "game the system" to increase the average test scores of students by, for example, excluding potentially low-performing students from the pool of students taking high-stakes tests, see Kim (2004) and Jacob (2002).

8. The effect of subgroup size on accountability measures was pursued in earlier work by Kane and Staiger (2001).

9. In addition, we analyzed data from the previous school year, 2000-2001; the results were qualitatively the same, and therefore we report results for only the most recent year for which data were made available.

10. Forty percent was a natural division in the data, allowing us to include schools with significant, but not predominant, special education populations while removing from the analysis schools that require separate, focused studies of their specific goals, operations, student and teacher populations, and instructional practices.

11. Grade size correlates very highly with school size in New York State.

12. The category non-White includes four of the five racial subgroups reported by the NYSED, including Hispanic, Black, Asian, and American Indian (White is the omitted group).

13. The same trends can be seen in the results from the previous year, 2000-2001. The major difference across years was the large increase in eighth grade math scores. Student pass rates increased by more than 8 points overall from 2000-2001 to 2001-2002. This increase was due to increases in all subgroups, including all disaggregated race groups: White pass rates increased by nearly 10 points; non-White pass rates increased by 7 points, as did Black and Hispanic pass rates; and Asian and American Indian pass rates increased by 8 points. These same increases were not seen in the eighth grade ELA exam results or in the fourth grade ELA and math exam results. Because all subgroups saw relatively large increases in eighth grade math pass rates, the size of the test score gap between White and non-White students did not change much over years.

14. In principle, the threshold should be high enough that the statistics are precisely measured but not so high that few schools will meet the threshold. Exact sizes are left up to states, however, and there is recent evidence that these sizes vary greatly for subgroups of disabled students, leading to wide state variation in numbers of schools failing to meet NCLB standards because of subgroup performance (Olson, 2005).

15. Note that New York State's approved accountability plan with the U.S. Department of Education shows that the state is still considering subgroup sizes between 20 and 40 (New York State Education Department, 2004; U.S. Department of Education, 2004a). Because the final minimum has yet to be determined and made public, and to be conservative, we chose the average of these two proposed cutoffs—30 across three grades, or 10 students per grade—which also reflects some of the most conservative minimums proposed by other state accountability plans. The lowest minimum (i.e., the outlier) subgroup size proposed for holding a school accountable is 5 in Maryland (with a 95% confidence interval). The next lowest minimums are 10 per grade in Louisiana, South Dakota, and Utah (with a 99% confidence interval); 11 in New Hampshire (with a 95% confidence interval); and 20 in New Jersey (with a 95% confidence interval) (see Olson, 2003).

16. The dissimilarity index can be interpreted as the proportion of Group A (e.g., non-White students) that would have to transfer from one school in which they are overrepresented to one in which they are underrepresented to achieve a distribution in each school proportional to their statewide percentage representation. The dissimilarity index is calculated according to the following formula:

$$D = 1/2 \sum_i \left| \frac{N_{Ai}}{N_A} - \frac{N_{Bi}}{N_B} \right|$$

where D is the dissimilarity index, N_{Ai} is the number of students from Group A (non-White students, in this case) in the i th school, and N_{Bi} is the number of students from Group B (White

students, in this case) in the i th school (see Carrington & Troske, 1997; Cortese, Falk, & Cohen, 1976; Massey & Denton, 1993; Zoloth, 1976).

17. Although a school may be held accountable for only White students, it may still contain some non-White students.

18. Similar trends were observed over years as well.

19. These numbers were similar for the previous year. In 2000-2001, 254 of 663 districts (38.3%) were accountable for both White and non-White fourth grade students, whereas 246 of 634 districts (38.8%) were accountable for eighth grade White and non-White students.

20. Regression results are available on request.

References

- Carnoy, M., & Loeb, S. (2003). Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis*, 25(1), 305-331.
- Carrington, W. J., & Troske, K. R. (1997). On measuring segregation in samples with small units. *Journal of Business and Economic Statistics*, 15(4), 402-409.
- Clotfelter, C. (2001). Are Whites still fleeing? Racial patterns and enrollment shifts in urban public schools, 1987-1996. *Journal of Policy Analysis and Management*, 20(2), 199-221.
- Cook, M. D., & Evans, W. N. (2000). Families or schools? Explaining the convergence in White and Black academic performance. *Journal of Labor Economics*, 18(4), 729-754.
- Cortese, C. F., Falk, R. F., & Cohen, J. K. (1976). Further considerations on the methodological analysis of segregation indices. *American Sociological Review*, 41(4), 630-637.
- Freedman, S. (2005, September 28). The achievement gap in elite schools. *The New York Times*, p. B7.
- Fryer, R. G., & Levitt, S. D. (2004). Understanding the Black-White test score gap in the first two years of school. *Review of Economics and Statistics*, 86(2), 447-464.
- Fuller, E., & Johnson, J. (2001). Can state accountability systems drive improvements in school performance for children of color and children from low-income homes? *Education and Urban Society*, 33(3), 260-283.
- Hanushek, E. (2002). *Publicly-provided education* (Working Paper 8799). Cambridge, MA: National Bureau of Economic Research. Retrieved February 25, 2003, from <http://dls.nber.org/papers/w8799.pdf>
- Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297-327.
- Hedges, L., & Nowell, A. (1998). Black-White test score convergence since 1965. In C. Jencks & M. Philips (Eds.), *The Black-White test score gap* (pp. 149-181). Washington, DC: Brookings Institution.
- Jacob, B. (2002). *Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools* (Working Paper 8968). Cambridge, MA: National Bureau of Economic Research. Retrieved May 11, 2004, from <http://papers.nber.org/papers/w8968>
- Jencks, C., & Philips, M. (1998). *The Black-White test score gap*. Washington, DC: Brookings Institution.
- Kane, T., & Staiger, D. (2001). *Improving school accountability measures* (Working Paper 8156). Cambridge, MA: National Bureau of Economic Research. Retrieved December 17, 2003, from <http://papers.nber.org/papers/w8156.pdf>
- Kane, T., & Staiger, D. (2003). Unintended consequences of racial subgroup rules. In P. E. Peterson & M. R. West (Eds.), *No child left behind? The politics and practice of accountability* (pp. 152-176). Washington, DC: Brookings Institution.

- Kim, D. Y. (2004, March). *The unintended consequences of New York State's high-stakes testing: Evidence from New York City public schools*. Paper presented at the 29th annual meeting of the American Education Finance Association, Salt Lake City, UT.
- Ladd, H. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review*, 18(1), 1-16.
- Linn, R. (2005). *Fixing the NCLB accountability system* (CRESST Policy Brief 8). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing. Retrieved October 4, 2005, from <http://www.cse.ucla.edu/products/newsletters/policybrief8.pdf>
- Loeb, S., & Strunk, K. (2003). The contribution of administrative and experimental data to education policy research. *National Tax Journal*, 56(2), 415-438.
- Massey, D., & Denton, N. (1993). *American apartheid*. Cambridge, MA: Harvard University Press.
- National Center for Education Statistics. (2003a). *The nation's report card: Mathematics highlights 2003*. Washington, DC: U.S. Department of Education. Retrieved December 17, 2003, from <http://nces.ed.gov/nationsreportcard/pdf/main2003/2004451.pdf>
- National Center for Education Statistics. (2003b). *The nation's report card: Reading highlights 2003*. Washington, DC: U.S. Department of Education. Retrieved December 17, 2003, from <http://nces.ed.gov/nationsreportcard/pdf/main2003/2004452.pdf>
- New York State Education Department. (2002). New York State school report card. Albany: Author. Retrieved July 8, 2002, from <http://www.emsc.nysed.gov/repercd2002/database.html>
- New York State Education Department. (2004). *Accountability peer review, New York State*. Retrieved March 8, 2004, from <http://www.emsc.nysed.gov/deputy/nclb/accountability/nclbaccountabilityplan.doc>
- Olson, L. (2003, August 6). "Approved" is relative term for Ed. Dept. *Education Week*, p. 1.
- Olson, L. (2005, September 21). AYP rules miss many in spec. ed.; more students left out of accountability ratings. *Education Week*, pp. 1, 24.
- Orfield, G., & Lee, C. (2004). *Brown at 50: King's dream or Plessy's nightmare?* Cambridge, MA: Harvard University, Harvard Civil Rights Project. Retrieved March 3, 2004, from <http://www.civilrightsproject.harvard.edu/research/reseg04/brown50.pdf>
- Paige, R. (2003, October 27). Real education reform [Letter to the editor]. *The New York Times*, p. A20.
- Persky, J., & Kurban, H. (2003). Do federal spending and tax policies build cities or promote sprawl? *Regional Science and Urban Economics*, 33, 361-378.
- Peterson, P., & West, M. R. (Eds.). (2003). *No child left behind? The politics and practice of accountability*. Washington, DC: Brookings Institution.
- Reardon, S. F. (2003). *Sources of educational inequality: The growth of racial/ethnic and socioeconomic test score gaps in kindergarten and first grade* (Working Paper 03-05R). State College: Pennsylvania State University, Population Research Institute. Retrieved January 3, 2005, from http://www.pop.psu.edu/general/pubs/working_papers/psu-pri/wp0305R.pdf
- Smith, S., & Mickelson, R. (2000). All that glitters is not gold: School reform in Charlotte-Mecklenburg. *Educational Evaluation and Policy Analysis*, 22(2), 101-127.
- Stiefel, L., Schwartz, A. E., & Chellman, C. C. (2003). *Education Finance Research Consortium condition report: Beyond Black and White: Patterns and relationships in New York State test score gaps*. Albany: Rockefeller College, State University of New York at Albany, Center for Policy Research.
- Stiefel, L., Schwartz, A. E., & Ellen, I. G. (2004, March). *Disentangling the race gap in students achievement*. Paper presented at the 29th annual meeting of the American Education Finance Association, Salt Lake City, UT.

- U.S. Department of Education. (2004a). *Accountability peer review, New York State*. Retrieved March 8, 2004, from <http://www.ed.gov/admins/lead/account/stateplans03/nycsa.doc>
- U.S. Department of Education (2004b, January 14). *Charting the course: States decide major provisions under No Child Left Behind*. Retrieved September 28, 2005, from <http://www.ed.gov/news/pressreleases/2004/01/01142004.html>
- Voith, R., & Gyourko, J. (2002). Capitalization of federal taxes, the relative price of housing, and urban form: Density and sorting effects. *Regional Science and Urban Economics*, 32, 673-690.
- Zoloth, B. S. (1976). Alternative measures of school segregation. *Land Economics*, 52(3), 278-291.

Leanna Stiefel (PhD, University of Wisconsin–Madison), is a professor of economics in New York University's Wagner Graduate School of Public Service, where she teaches courses in multiple regression, the economics of education, and microeconomics. Her current research includes studies of the measurement of school efficiency, the effect of school size on costs and student performance, the achievement of immigrant students, test score gaps within schools, and the effects of school organization on student outcomes. She is the author of *Statistical Analysis for Public and Nonprofit Managers* (Praeger, 1990), a coauthor (with Robert Berne) of *The Measurement of Equity in School Finance: Conceptual, Methodological, and Empirical Dimensions* (Johns Hopkins University Press, 1984), and a coeditor (with R. Rubenstein, A. E. Schwartz, and J. Zabel) of *Measuring School Performance and Efficiency: Implications for Practice and Research* (Eye on Education, 2005). Her work is published in *Education Evaluation and Policy Analysis*, the *Journal of Human Resources*, the *Journal of Policy Analysis and Management*, and *Education Finance and Policy*, among other journals. She is a past president of the American Education Finance Association, a member of the National Center of Education Statistics Technical Planning Panel, on the policy council of the Association for Public Policy Analysis and Management, and a governor on the New York State Education Finance Research Consortium.

Amy Ellen Schwartz (PhD, Columbia University) is director of the Institute for Education and Social Policy and a professor of public policy, education, and economics at the Steinhart School of Education and Wagner School of Public Service at New York University. Her current research reflects her interest in urban education policy and disadvantaged students, considering whole school reform, resource equity, accountability systems, and the creation of small schools. Her work has been published in *Educational Evaluation and Policy Analysis*, the *Journal of Human Resources*, and *Education Finance and Policy*, among other academic journals. The author of several book chapters, she coedited the *2005 Yearbook of the American Education Finance Association: Measuring School Performance and Efficiency*. Her current work examines the education of immigrant students (funded by the Spencer Foundation and the Rockefeller Foundation), the organization of primary schooling (funded by the U.S. Department of Education's Institute of Education Sciences), and the cost of small schools (funded by the Carnegie Corporation).

Colin C. Chellman is a research associate at the Institute for Education and Social Policy at New York University, working on issues of urban policy, education finance, and equity. He is a graduate of the master's of public administration program at New York University's Wagner Graduate School, where he specialized in public finance and financial management. He received his BA in international politics from Johns Hopkins University and is currently a PhD student in public finance, financial management, and urban policy at New York University's Wagner School.