728

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208269
27 Hillhouse Avenue
New Haven, CT 06520-8269

CENTER DISCUSSION PAPER NO. 728

INTRASCHOOL VARIATION IN CLASS SIZE: PATTERNS AND IMPLICATIONS

Michael Boozer
Yale University

Cecilia Rouse
Princeton University
and NBER

May 1995

Note: Center Discussion Papers are preliminary materials circulated to stimulate discussion and critical comments.

Financial support was provided by the Center for Economic Policy Studies and the Industrial Relations Section, both at Princeton University. Cecilia Rouse thanks the Russell Sage Foundation and the Spencer Foundation for their financial support.

For helpful comments, we thank Kristin Butcher, David Card, Earl Kim, Alan Krueger, Steve Pischke, Paul Schultz, members of the Columbia University Labor Seminar and Princeton Labor Lunch, and participants at the NBER Conference on the Well-being of Children. We also thank Herb Abelson, Sybil Chabane, Danielle Gordon, Elana Hahn, David Lee, Karen Livescu, Kerry Radvany, Kevin Radvany, Dirk Reinhardt, Mark Reinhardt and Tara Yasso for expert help with the survey.

ABSTRACT

Economists attempting to explain the widening of the black-white wage gap in the late 1970's by differences in school quality have been faced with the problem that recent data reveal virtually no gap in the quality of schools attended by blacks and whites using a variety of measures. In this paper, we re-examine racial differences in school quality. We begin by considering the effects of using the pupil-teacher ratio, rather than the school's average class size, in an education production function since the pupil-teacher ratio is a rough proxy, at best. Second, we consider the importance of using actual class size rather than school-level measures of class size.

We find that while the pupil-teacher ratio and average class size are correlated, the pupil-teacher ratio is systematically less than or equal to the average class size. Mathematically, part of the difference is due to the intraschool allocation of teachers to classes. As a result, while the pupil-teacher ratio suggests no black-white differences in class size, measures of the school's average class size suggest that blacks are in larger classes. Further, the two measures result in differing estimates of the importance of class size in an education production function. We also conclude that school level measures may obscure important within-school variation in class size due to the small class sizes for compensatory education. Since black students are more likely to be assigned to compensatory education classes, a kind of aggregation bias results. We find that not only are blacks in schools with larger average class sizes, but they are also in larger classes within schools, conditional on class type. The intraschool class size patterns suggest that using within-school variation in education production functions is not a perfect solution to aggregation problems because of non-random assignment of students to classes of differing sizes. However, once the selection problem has been addressed, it appears that smaller classes at the eighth grade lead to larger test score gains from eighth to tenth grade and that differences in class size can explain approximately 15 percent of the black-white difference in educational achievement.

KEY WORDS: Economics of Education, Human Capital

I. Introduction

For decades before the 1970's the difference in the average wages for black and white workers had been decreasing, but in the mid-1970's this trend slowed and then reversed. Authors such as Smith and Welch (1989), and Juhn, Murphy, and Pierce (1991) have suggested that the lower quality of schooling for black students relative to white students may account for lower acquired skills of black workers. If the price of skills then increases, the gap in wages for blacks and whites will widen. However, most national data, such as the Common Core of Data Surveys, the High School and Beyond, and the National Longitudinal Study of the High School Class of 1972, report no discernable difference in the quality of schools attended by blacks and whites.[1] For example, in the Common Core the mean pupil-teacher ratio in schools attended by blacks, whites, and hispanics is 18.16, 18.36, and 20.33, respectively (Boozer, Krueger, Wolkon (1992))[2]. It is only in more "exotic" measures of school quality, such as levels of computer usage, that differences between the races begin to appear.

In this paper, we re-examine racial differences in school quality. In order for school quality to explain the black-white gap in achievement two relationships must hold: black students must attend schools of inferior school quality, and school quality must matter for achievement. We look for this pattern and implication in two ways. First, we consider the effects of using the pupil-teacher ratio, rather than the school's average class size, in an education production function since most researchers, lacking data on actual class size, use the pupil-teacher ratio, and because most admit that it is a rough proxy, at best. Second, we consider the importance of using actual class size rather than school-level measures of class size in an education production function.

---

[1] For example, see Boozer (1992), Boozer, Krueger, Wolkon (1992), and Grogger (1994). It is important to note that we only address one component of school quality, namely class size. These papers also report there are no discernable differences in other measures, such as teacher quality.

[2] The larger pupil-teacher ratio for hispanics is largely attributable to the fact that they reside in the West where class sizes are larger.

We find that while the pupil-teacher ratio and average class size are correlated, the pupil-teacher ratio is systematically less than or equal to the average class size. Mathematically, part of the difference is due to the intraschool allocation of teachers to classes. As a result, while the pupil-teacher ratio suggests no black-white differences in class size, measures of the school's average class size indicate that blacks are in larger classes. Further, the two measures result in differing estimates of the importance of class size in an education production function. We also find that the pupil-teacher ratio obscures important within-school variation in class size due to the small class sizes for compensatory education.[3] Since black students are more likely to be assigned to compensatory education classes, a kind of aggregation bias results. We find that black students are in schools with larger average class sizes and are in larger classes within schools, as well, conditional on class type.

Our results have important implications for estimating education production functions in general, and black-white gaps in achievement or labor market outcomes in particular. Within-school class size patterns indicate that using within-school variation in production functions is not a perfect solution to aggregation problems because of non-random assignment of students to classes of differing sizes. However, once the selection problem has been addressed, it appears that smaller classes in 8th grade lead to larger test score gains and that differences in class size can potentially explain approximately 15 percent of the black-white difference in the gain in test score between 8th and 10th grade. Our results also suggest that a class size intervention at a later grade would have a smaller effect, at least in terms of black-white differences, a result perhaps due to relatively greater drop out activity by black students relative to white

---

[3] In this paper, we use the term "compensatory education" to refer to special (or special needs), remedial, and bilingual education. "Special Needs" refers to classes for students with disabilities. The categories include students who are: hard of hearing, mentally retarded, multihandicapped, orthopedically and/or other health impaired, seriously emotionally disturbed, specific learning disability, speech impaired, and visually handicapped (Staff to Student Ratios: Class Size/Caseload, 1986). Remedial education is designed to compensate for past economic or environmental deficiencies, and bilingual education to facilitate the learning of English.

students. And, we find that a smaller class size at 8th or 10th grade would have no discernable impact on subsequent dropout rates from high school.

The next section of the paper describes the data, sections three and four consider the patterns and implications of using the school's average class size or actual class size in education production functions, section five considers the implications for the black-white difference in school achievement, and section six concludes.

II. The Data

In order to examine the importance of intraschool variation in classes, we need data that measure the overall pupil-teacher ratio, average class size, and individual class sizes within schools. We therefore rely on two data sets: a survey of teachers that we conducted in New Jersey and the National Longitudinal Survey of 1988 (NELS).

The New Jersey Survey of Teachers

Much of our descriptive data come from a telephone survey of a random sample of 500 teachers in NJ that we conducted in June, 1994. We asked the teachers about the sizes of their classes, the average class size in the school, the numbers of students and teachers in the school, and the racial compositions of both their classes and of the school.[4] We also asked the teachers for the name, district, and county of their school in order to merge these data with data from the NJ Department of Education and the Common Core. In all, the surveyed teachers and their schools appear representative of teachers and schools in the state. (See Tables 1 and 2 in the Data Appendix for the mean characteristics of teachers in the NJ Survey and all classroom teachers in New Jersey and the mean characteristics of the schools in the NJ Survey compared to the mean characteristics of all schools in New Jersey.)

---

[4] The survey instrument is available from the authors upon request.

We merge these data with data from the Common Core of Data Surveys and the New Jersey Department of Education. The Common Core is a national survey of schools that provides information on the racial composition of most schools in the U.S., as well as the student enrollment figures and the number of full-time equivalent instructional staff (FTE's) from which we can construct a pupil-teacher ratio. The data from the NJ Department of Education's administrative records are individual data on all certificated teachers in New Jersey.[5] The data identify each teacher's school and class subject. We also have the total enrollment and the racial composition for each school in New Jersey as of October, 1992.

The National Education Longitudinal Survey of 1988 (NELS)

The NELS is a national, stratified sample of eighth-graders in 1988 who were followed-up in 1990 and 1992. In addition to surveying the students, the U.S. Department of Education also surveyed school administrators (creating the schools survey) and two teachers (each of whom taught from one of the four main academic subject areas: mathematics, science, reading, and history) for each student in the survey. These teachers were asked questions about the student's class, including the class size. Because approximately 20 students were sampled within each school, we can use these data to estimate school level averages for measures not provided by the school survey. Further, since students were given tests in each year in history, math, reading, and science, we can create student-class records in which we match the class size with the test score gain for that particular subject.

[5] The data are from the 1992-1993 school year, the most recent year for which we have been able to obtain data. We thank Howard Bookin of the NJ Department of Education for making these data available to us.

We use two basic samples.[6] To study the first follow-up (10th grade) outcomes, we use a sample that includes public school students with complete data who participated in both the base year and first follow-up surveys, resulting in 20,131 observations which comprises 10,499 individuals from 751 schools. To model the second follow-up (12th grade) outcomes, we include public school students who participated in all three waves resulting in 10,369 observations comprised of 6,692 students from 698 schools. We use the appropriate first or second follow-up panel weights supplied on the NELS data. The means of our basic samples are in Appendix Table 1.[7]

III. The Pupil-Teacher Ratio vs. Average Class Size

We begin by considering whether previous researchers have found no difference in black and white school quality because they have used the pupil-teacher ratio, rather than the school's average class size in the education production function. Figures 1a and 1b illustrate that the pupil-teacher ratio and the school's average class size, while correlated, may reflect different aspects of the school's teaching resources.[8] Figure 1a graphs the average class size and the pupil-teacher ratio against the percentage of

---

[6] We use slightly different samples when studying drop-out behavior since students who had dropped-out were disproportionately likely to be missing class sizes and test scores, and since we averaged the class sizes and test scores resulting in only one observation per student. The means of these samples are available from the authors upon request.

[7] The sample size as of the second follow-up is about one-half the size as of the first follow-up, because we do not require that all students in the first follow-up also be participants in the second. Further, not all students continued to be enrolled in the same (two) subjects in the 10th grade as they were in the 8th grade, a requirement for our second follow-up analysis. Comparing the (appropriately) weighted means of region, race, and family income variables in the two follow-ups to the (weighted) initial base year sample reveals that the panel weights aid in reducing the non-comparability of results from the different follow-ups. That said, when we restrict the sample used in the analysis of test score outcomes in the first follow-up to members in the second follow-up, we find a different result in one of the four cases, as we report in Appendix Table 3. Thus, we are somewhat worried about the large amount of attrition from the sample as of the second follow-up, and tend to have relatively more confidence in the first follow-up results than the second follow-up results.

[8] The correlation between the pupil-teacher ratio and the average class size is relatively low at 0.13 in the NJ Survey, and 0.26 in the NELS.

FIGURE 1A

Comparison of Average Class Size

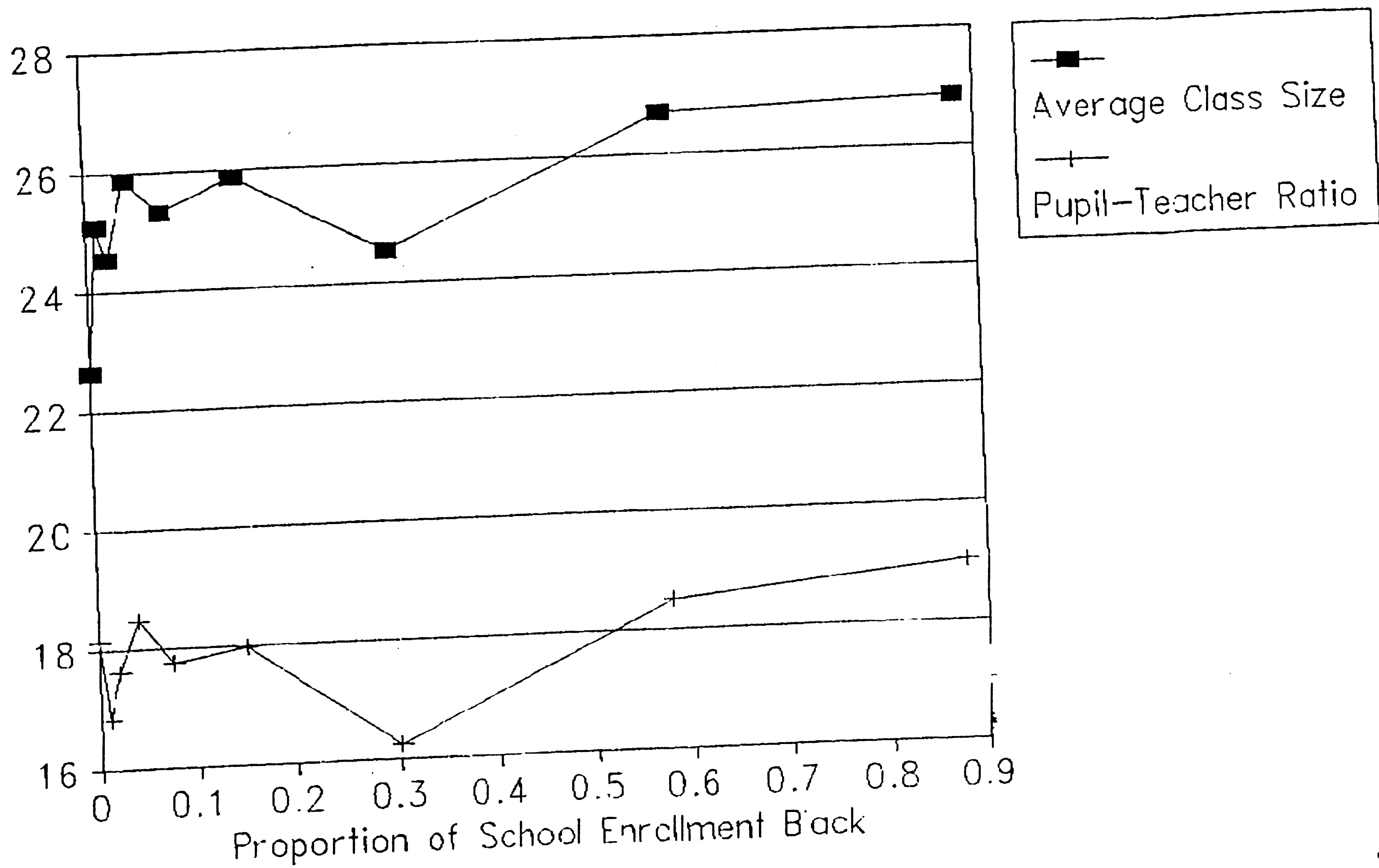
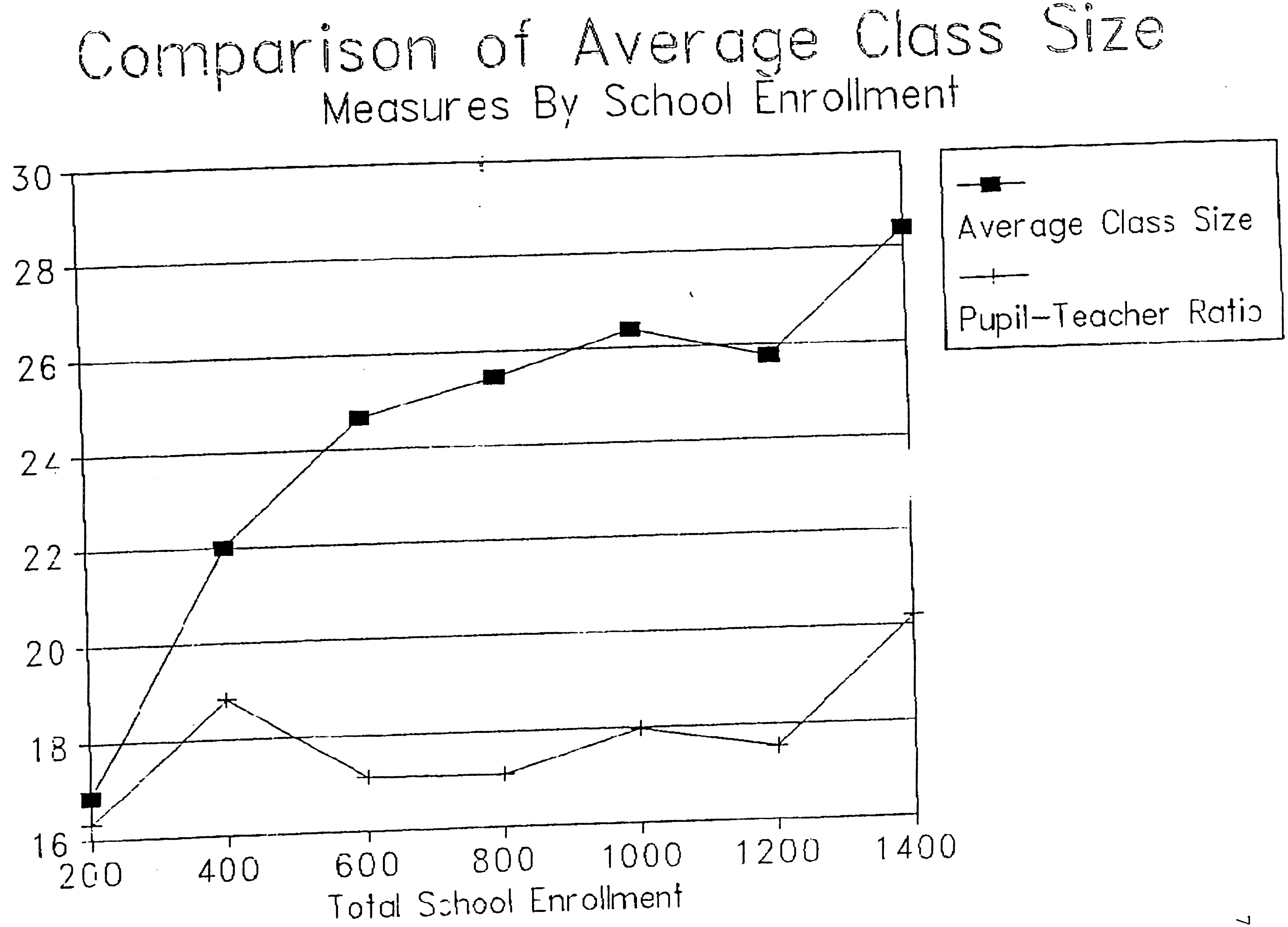


FIGURE 1B


Comparison of Average Class Size
Measures By School Enrollment
30
28
26
24
22
20
18
16
200
400
600
800
1000
1200
1400
Total School Enrollment
Average Class Size
Pupil-Teacher Ratio

the school's enrollment that is black using the NELS data. Figure 1b graphs the average class size and the pupil-teacher ratio against the school size, also using the NELS. In Figure 1a, the pupil-teacher ratio is uniformly smaller than the average class size and the lines diverge slightly as the percentage of students in the school increases.[9] Note that the divergence appears primarily driven by the extremes in the racial composition distribution. The two measures widen as school size increases in Figure 1b.

One potential explanation for these patterns comes from the mechanical relationship between the two measures.[10] One can describe the relationship by taking a Taylor series expansion of average class size around its mean, and then taking expectations of both sides to get:

$$E(CS)_s = PT_s\{ 1 + \frac{\sigma^2(T)_s}{(E(T)_s)^2} + R \} \qquad (1)$$

where $E(CS)_s$ is the average class size for school s, PT is the pupil-teacher ratio of school s, $\sigma^2(T)_s$ is the variance of the number of teacher per class within school s, T is the number of teachers assigned to each class, and R is a remainder term.

The average class size will be greater than or equal to the pupil-teacher ratio, and the two measures will diverge as the variance in how teachers are distributed across classes within schools increases. If there is only one teacher per class in a school, the pupil-teacher ratio will equal the class size, if only classroom teachers are included in the pupil-teacher ratio. On the other hand, if some of the

---

[9] Other researchers having access to both pupil-teacher ratios and average class sizes have noted that the average class size is larger than the pupil-teacher ratio (Blinder (forthcoming), Flyer and Rosen (1994)). For example, in the NELS the average class size is 24 compared to an average pupil-teacher ratio of 18. However, this, alone, need not mean that the pupil-teacher ratio captures a different dimension of school quality than the average class size. Even if the pupil-teacher ratio is constructed with only classroom teachers, the average class size will be greater than or equal to the pupil-teacher ratio by Jensen's Inequality; see equation (1).

[10] It is also possible that measures of the pupil-teacher ratio in national data mis-represent average class size because either the number of students or teachers are mis-counted. We have explored this possibility to some extent and found evidence that the number of teachers may be inflated, perhaps due to the inclusion of "non-regular" teachers among the teaching staff.

classes are team taught and others are not, the average class size will diverge from the pupil-teacher ratio. One factor that contributes to such variation is the number of remedial and special education classes since they are more likely to be taught by more than one teacher.[11] If classes are more likely to be team taught in predominately black schools, thereby increasing the variance of teachers within the schools, then the pupil-teacher ratio will diverge from the average class size more in black schools than in white schools.[12]

Patterns and Implications

To determine if the pupil-teacher ratio gives a different description of school class size than does the average class size, we regressed the pupil-teacher ratio and average class size on the racial composition of the school using the three data sets on schools in New Jersey and the NELS data for base year schools in the northeast only.[13] The results are in Table 1. Regression coefficients using the Common Core, the NJ Department of Education, and the NELS data suggest that blacks and hispanics are in schools with marginally smaller pupil-teacher ratios, although these differences are not statistically significant at conventional levels. In the NJ Survey, we find that blacks and hispanics are in schools with marginally larger pupil-teacher ratios, although again the differences are not significant. Thus, according to the pupil-teacher ratio, blacks and hispanics do not attend schools of inferior quality. The conclusion reverses for the average class size. In both surveys, blacks attend schools with larger average class sizes. The magnitudes of the coefficients are remarkably similar between the two surveys and the coefficient estimates are statistically significant. The results for hispanics are mixed. In the NJ Survey, it appears

[11] Authors' calculations from the NJ Survey.

[12] It is not clear which measure better reflects the resources of the school. A class of 30 students with two teachers may not be so different from a class with 15 students and one teacher.

[13] For the school level analyses we only use the base-year school information since the first and second follow-up schools are not a random sample of schools (National Education Longitudinal Study of 1988, Second Follow-Up: School Component Data File User's Manual (1994)).

Table 1

Regression of Pupil-Teacher Ratio and Class Size on the Racial Composition of the Student Body: A Comparison of Common Core, NJ Dept. of Education, the NJ Survey, and NELS Data

Data Set:	Common Core	NJ Dept. of Education	NJ Survey		NELS (Northeast Only)[a]	
Dependent Variable:	Pupil-Teacher Ratio	Pupil-Teacher Ratio	Pupil-Teacher Ratio	Avg. Class Size	Pupil-Teacher Ratio	Avg. Class Size
% Black	-0.529 (0.853)	-1.232 (0.818)	2.884 (1.927)	2.935 (0.835)	-0.935 (1.022)	3.855 (1.583)
% Hispanic	-0.830 (1.054)	-0.214 (1.011)	2.821 (2.381)	4.463 (1.031)	-2.973 (1.743)	0.981 (2.701)
Constant	15.981 (0.235)	15.330 (0.226)	17.961 (0.532)	22.460 (0.231)	13.875 (0.274)	21.595 (0.425)
Mean of Dependent Variable [Std. Dev.]	15.820 [3.292]	15.131 [3.166]	18.665 [7.474]	23.340 [3.397]	13.577 [2.635]	22.156 [4.117]
R^2	0.004	0.008	0.013	0.104	0.037	0.052
No. of Observations	326	326	326	326	128	128

Notes: Standard errors are in parentheses.

[a] These regressions are weighted using the base year school weights.

that hispanics attend schools with larger classes while in the NELS the coefficient is positive, but insignificant.

We further assess the effects of using the pupil-teacher ratio rather than the average class size in an education production function using all base year schools in the NELS in Table 2. These regressions use school-level data and control for school characteristics and average family background characteristics of the students. Once again, we find that the pupil-teacher ratio does not (statistically) increase in schools with a larger proportion of black students, but that the average class size does. In addition, the pupil-teacher ratio suggests that hispanic students are in schools with significantly smaller pupil-teacher ratios while the gap is smaller in magnitude and statistically insignificant in the average class size.[14] The second two columns assess the effect of using the pupil-teacher ratio instead of the average class size in an education production function. Here we regress the average 8th grade to 10th grade test score gain by students in the school on school and family background characteristics. We find that the pupil-teacher ratio has essentially no effect on the test score gains of students. On the other hand, students in schools with larger average class sizes have significantly smaller test score gains.[15]

Boozer, Krueger, and Wolkon (1992) and Grogger (1994) use the pupil-teacher ratio and report that, if anything, black students attend schools with smaller average pupil-teacher ratios than white students; Grogger (1994) therefore concludes that school quality cannot explain the recent divergence in the black-white wage gap. However, we find that the pupil-teacher ratio and average class size have differing patterns (at least in recent data) that can generate contrasting descriptions across schools and can

---

[14] The coefficient on percent hispanic is negative in this table because we control for the junior high school's region of the country.

[15] This result is similar to that found by Blinder (forthcoming) who used the average class size of the school as an instrument for the actual class size of the school. Since this procedure sweeps out the within school source of variation in class size, her IV procedure is very similar to doing OLS across schools.

Table 2

Pupil-Teacher Ratio vs. Average Class Size:
Descriptive and Production Function Differences
(NELS)

	Dependent Variable			
Independent Variables	Pupil-Teacher Ratio	Average Class Size	Avg. Test Score Gain	Avg. Test Score Gain
Pupil-Teacher Ratio			0.017 (0.019)	
Average Class Size				-0.048 (0.017)
Percent Black	0.418 (0.761)	2.086 (0.879)	0.021 (0.399)	0.126 (0.398)
Percent Hispanic	-3.623 (0.883)	-1.246 (1.021)	0.349 (0.468)	0.220 (0.461)
R^2	0.244	0.236	0.039	0.049
No. of Observations	760	760	748	748

Notes: Standard errors in parentheses. All regressions are weighted using the base year school weight. Other regressors include a constant, the average family income, the average family income squared, the proportion of students from homes with a computer and who live with both parents, the total enrollment in the school, urban and rural dummies, and region dummies. The average test score gain is between the 8th and 10th grades.

also affect education production function estimates.[16] Thus, using only the cross-school variation in class size we find that it satisfies the two requirements for explaining the black-white gap in achievement: black students are in larger classes than are white students, and smaller class sizes are associated with larger test score gains. In the next section we address the importance of the within school variation in class size in describing the racial patterns and the selection problem it creates for measuring the implications.

IV. The Role of Class Type: Aggregation Bias

The fact that school average class size matters, but pupil-teacher does not suggests a second reason that researchers may not have detected racial differences in class size: the pupil-teacher ratio (more than the average class size) may obscure important intraschool variation in class size. To see this, let CS_{is} be the class size of student i in school s (assuming for the moment only one class size for each student to keep notation simple), and let $\overline{C}_s$ be the (sample) average class size in the school. The question is whether the regressor $\overline{C}_s$ (or its proxy, the pupil-teacher ratio) captures variation in school quality that is important in the education production function. If $\overline{C}_s$ is a noisy measure of CS_{is} such that

$$\overline{C}_s = CS_{is} + \varepsilon_{is}$$

where ε_{is} is an error term which is negatively correlated with CS_{is}, then using $\overline{C}_s$ as a regressor in a schooling production function introduces downward bias in the estimated coefficient on class size. In this case, using school-level measures would not bias estimates of, for example, the black-white difference in educational outcomes. On the other hand, if ε_{is} is correlated with other explanatory variables, then using school-level measures will generate misleading production function estimates. For example, if black

[16] In contrast, Card and Krueger (1992) use state-level measures of pupil-teacher ratios and find that blacks were historically in schools with larger classes. While we can only speculate, part of the explanation for why this pattern does not appear in more recent data may be that during the period in which they measured school quality (the early part of this century), the average class size and pupil-teacher ratio were more closely correlated. This might be true if, for instance, the intraschool allocation of teachers did not vary widely across schools which the growth in compensatory education programs (since the 1960s) may have generated.

students are assigned to larger classes within schools, and whites to relatively smaller classes, then a study of the black-white gap in school quality using only the cross-school variation will tend to understate the magnitude of that gap.

One potential source of error is the distribution within schools of gifted, remedial, and special education classes. If the number, or proportion, of students classified as remedial (or special education) across schools varies (i.e. so it is not always the bottom one-third of students in the test score distribution), then schools with a higher proportion of remedial classes will have smaller average class sizes. In addition, school-level measures may aggregate across heterogeneous production processes (Hanushek (1979)). It is possible that remedial or special education classes produce test score gains with a different production technology than above-average or gifted classes. If remedial and special education classes have smaller class sizes and generate lower test score gains for a given class size than do high achieving classes, then the fitted regression line that ignores these differences will estimate an upward sloping relationship between class size and test score gains. This presents a serious problem for estimating education production functions.

Theoretically, if one has within-school class sizes, then a regression including school fixed effects and allowing different slopes for the remedial and high-achieving classes should recover the hypothesized negatively sloped relationship between test score gains and class size. However, even within these classifications, students are likely allocated to classes according to ability so that one must also account for the allocation mechanism.

A. Patterns

Table 3 assesses the extent to which this "aggregation bias" obscures intraschool differences in class sizes. In the upper panel, we report the racial composition across class types (the rows sum to 100) using the NJ Survey. In the lower panel, we present the average class size by race and type of class.

Table 3

Racial Composition by Class Type[a]

Type of Class	Distribution (percentage)	% Black	% Hispanic	% White	% Other
Regular	72.7	13.9	8.7	71.1	5.7
Special Needs	14.2	22.9	19.3	72.0	3.6
Remedial	2.6	4.3	27.9	61.8	6.1
Gifted	4.7	8.3	3.7	68.1	14.9
Other	5.7	2.9	22.3	64.9	5.8
Total/Overall	100.0	14.1	11.2	70.5	5.9

[a] Each row indicates the race composition for that type of class. The rows may not sum to 100% since respondents were not constrained to make the racial composition of their classes sum to 100%.

Average Class Size by Race and Class Type

	Average Class Size			
Type of Class	Black[b]	Hispanic[b]	White[b]	Overall
Regular	24.4	24.4	22.6	22.1
Special Needs	11.8	13.8	12.1	9.7
Remedial	12.8	14.1	10.7	10.0
Gifted	24.6	28.3	22.7	22.2
Other	25.4	22.9	22.9	18.6

[b] Means are calculated by weighting the class size by the percentage of the class that is black, hispanic, or white.

Notes: Data from the New Jersey Survey of Teachers. There are 422 observations.

Although blacks represent 14% of the students overall, they represent 23% of the students enrolled in special needs classes. Hispanics, though 11% of students overall, comprise 28% of those in remedial classes. These racial differences have implications for the class sizes experienced by students since the average class size varies by type of class. For example, a "regular" (or average) class has 22 students on average, compared to 9.7 for students in special needs classes and 10 students in remedial classes. To the extent that black and hispanic students are disproportionately represented in these smaller classes, the average class size (and the pupil-teacher ratio), may not accurately reflect their schooling inputs, particularly for the "average" student.

Table 4 estimates the importance of the type of class on the class sizes of black and white students. In the first two columns, we use the NJ Survey and regress the teacher's class size on the percentage of the class that is black and hispanic, whether the teacher is at an elementary or junior high school, and whether the teacher teaches alone (as opposed to having teachers' aides or team teaching). There is no statistically significant difference in the class size at the level of the school and, in column (1), teachers who teach alone do not have substantially smaller classes.[17] And, there is little or no difference in the class sizes of blacks and hispanics.

In column (2) we condition on the type of class ("regular" classes are the omitted category). The large increase in the R^2 indicates that the type of class is a big determinant of the class size. Conditional on the type of class, teachers who teach alone have almost 2.5 fewer students, and remedial and special needs classes have 13 fewer students than regular classes. We find no difference between the class sizes of gifted and regular classes, however. The racial differences are now pronounced. In particular, as the

[17] Boozer, Krueger, Wolkon (1992) report that the pupil-teacher ratio in elementary and junior high schools is larger than in high schools. The negative signs on the coefficients for school type in Table 4 suggest just the opposite when using the average class size. Interestingly, if we regress the pupil-teacher ratio on the type of school using the NJ Survey we get results similar to Boozer, Krueger, Wolkon. The difference may be due to increased use of team teaching and teaching specialists in elementary schools.

Table 4

Class Size As Explained By Class Racial Composition and Type

Independent Variable	Data Set: NJ Survey		NELS	
Percent Black/ Student is Black[a]	1.475 (1.571)	3.287 (1.184)	1.125 (0.108)	1.647 (0.107)
Percent Hispanic/ Student is Hispanic[a]	-1.843 (1.529)	2.516 (1.164)	1.323 (0.125)	1.666 (0.123)
Elementary School	-0.353 (0.816)	-0.687 (0.614)		
Junior High School	-1.934 (1.217)	-1.617 (0.901)		
Teach Alone	0.192 (0.904)	-2.378 (0.684)		
Type of Class:				
Gifted		0.604 (1.226)		
Remedial		-12.539 (1.628)		
Special Needs		-13.490 (0.769)		
Other		-3.613 (1.125)		
Achievement Level of Class:				
High				0.617 (0.088)
Low				-2.867 (0.101)
R^2	0.011	0.466	0.009	0.051
No. of Obs.	422	422	22641	22641

Notes: Standard errors are in parentheses. All regressions also include a constant. The regressions using the NELS are weighted by the panel weights.

[a] Race in the NJ Survey is the percentage of the class that is black/hispanic; race in the NELS is the race of the student.

percentage of hispanic or black students increases by 10 percentage points, the class size increases by 0.25 to 0.3 students.

We performed a similar analysis using the base year NELS. The survey does not report the racial composition of each class, so we regress the student's class size on his or her race. Further, the teachers do not report whether the class is "remedial", "gifted", or "regular" although they do report their assessment of the overall achievement level of the class.[18] We find that blacks and hispanics are in slightly larger classes than non-minority students. Controlling for the achievement level of the class again increases the coefficients by almost 50% for black students and by 25% for hispanics. It is not surprising that controlling for the type of class in the NELS does not change the coefficients on race as dramatically as in the NJ Survey because the achievement level is an imprecise measure of the type of class. As evidence, the magnitudes of the coefficient estimates of "high" and "low" achievement are not as large as the coefficients on class type in the NJ Survey.

We decomposed the effect of the type of class on class size to understand whether it was the characteristics of students or the school policies that were generating the differences by race.[19] If black and hispanic students had the same distribution of characteristics as white students, they would be in schools with slightly larger average class sizes. More importantly, if black and hispanic students were in schools with the same class size patterns as white students, they would have lower average class sizes than white students. It appears that the patterns within the schools matters slightly more than the characteristics

[18] We get similar results when we use the more detailed descriptions of class types that exist for the 10th and 12th grade classes.

[19] We modeled the relationship between class size and type of class weighted by either the number of black, white, or hispanic students in the class using both the NJ Survey and the NELS. We then computed the mean class size using the coefficient estimates based on the white students and the means of the independent variables for the black or hispanic students, and then calculated the mean using the black or hispanic coefficient estimates and the means of the independent variables for the white students. The results are available from the authors upon request.

of the students. Thus, to some extent the seemingly similar class sizes between minority and white students at the school level is the result of class type.

B. Implications

Our results suggest that, conditional on class type, blacks are in larger classes than whites. However, we must still test whether actual class size matters in an education production function leaving an opportunity for school quality to account for part of the black-white gap in labor market outcomes. Unfortunately, given that class size is largely determined by the type of class, one cannot simply interpret OLS coefficients of actual class size on educational outcomes as causal. In particular, one must address the fact that students are not randomly sorted into classes of different sizes. In this section, we use within-school variation in class size to estimate education production functions. In section five, we then estimate the implied effect of class size on the black-white gap in educational attainment.

1. Modeling Education Production Functions

In the following analysis, we model the effect of class size on academic achievement as measured by test scores and the likelihood of dropping out.[20] Consider, first, a simple cross-sectional model,

$$T_{it} = a + X_i b + gC_{it-1} + e_{it} \tag{2}$$

where T_{it} is the test score of the student in period T, X_i is a vector of individual and school characteristics, C_{it-1} is the class size of the student in period t-1, and ε_{it} is an error term. We have in mind an experiment in which we "treat" a student at time t-1 with a class of a given size in the relevant subject for the

[20] See, also, Hanushek (1979) and Hanushek and Kain (1972) for discussions of the issues involved in choosing a functional form for an education production function.

duration of period t-1, and then observe the impact on his/her test score at time t.[21] In this specification, we assume that the class size in the previous period is all that is relevant for the current test score, or that the class size at time t-1 is a sufficient statistic for the entire past history of the educational production function since educational production processes are inherently cumulative.

However, it is more realistic to assume that the current school quality (class size) does not fully summarize the student's school quality history and that previous school quality also matters for current school achievement. In this case, if we do not control for past schooling quality then the effect of class size on test scores, if measured as late as the 8th grade, will be more reflective than causal, even if students were initially randomly assigned to classes. In this case, the education production function has the following form,

$$T_{it} = \alpha + X_i\beta + \gamma_1 C_{it-1} + \gamma_2 C_{it-2} + \epsilon_{it} \tag{3}$$

where current educational achievement is a function of last period's class size as well as previous class sizes (the assumption of only one extra lag is for expositional convenience only -- in reality we may take C_{it-2} to be a vector of the entire past history, or a scalar summary statistic that serves as its proxy). If we assume that C_{it-1} and C_{it-2} are correlated, so that $C_{it-1} = \rho\, C_{it-2} + v_{it-2}$, then plim $\hat{g} = [\gamma_1 + (\gamma_2/\rho)]$ (where $\hat{g}$ is from equation (2)). If class sizes are positively correlated over time (i.e., $\rho > 0$), and γ_1 and γ_2 have the same sign, then $\hat{g}$ will overstate the role of C_{it-1} in determining test scores, but should still give an indication of the sign of the effect of class size as well as an estimate of the cumulative effect of class

[21]In reality, however, we observe the data at two year intervals, although we suppress this feature in our discussion of the choice of appropriate functional form.

size on academic achievement, assuming, of course, that students are assigned to classes of varying sizes conditionally independent of the error term.[22]

One way in which we address the lack of data on the past history of schooling inputs is by estimating an equation of the form,

$$T_{it} = \alpha + X_i\beta + \gamma C_{it-1} + \delta T_{it-1} + \epsilon_{it} \qquad (4)$$

where we include T_{it-1} as a regressor in order to hold constant the determinants of the test score at time t-1, which would include the entire vector of past schooling inputs. We include T_{it-1} precisely because we lack a true experimental situation at the 8th grade. A constrained version of equation (4) sets $\delta = 1$, and involves regressing the test score gain on the class size level,

$$T_{it} - T_{it-1} = \alpha + X_i\beta + \gamma C_{it-1} + \epsilon_{it} \qquad (4')$$

which estimates the value-added from a given class size treatment administered at time t-1. In addition, since T_{it-1} is determined by the past history of school quality and individual attributes not included in the X's, specification (4′) partially controls for a person-specific fixed effect (in terms of the test score levels, but not the gains). Another way in which we address this problem is by estimating equation (3) for the 12th grade outcomes for which we partially observe the past history of class size inputs (from the 8th and 10th grades). We are somewhat agnostic as to whether the equations using the test score levels or the gain should be preferred, and therefore report results based upon both specifications.

---

[22] In the NELS, the correlation between the 8th and 10th grade class sizes is 0.27.

In the presence of a "true" randomized experiment, whereby students are assigned to classes of varying sizes, equations (3) and (4) (or (4′)) would clearly identify the impact of class size on school achievement. Without a true randomized experiment, however, we must assume that students are assigned to classes of varying sizes in a way that is uncorrelated with test scores at time t, conditional on the included X's. Thus, while OLS regression applied to equations (3) and (4) (or (4′)) will address the problem of omitted schooling histories, they will tend to yield positive estimates of g. However, these estimates are clearly not causal.[23] They are driven by the selection problem, noted above, that lower "ability" students tend to be assigned to smaller classes, on average, and these lower "ability" students also tend to score lower on achievement tests. As a result, one must employ either a true randomized experiment, or an instrumental variables strategy.

2. Evidence from the NELS: Using State Policy as an Instrument

Given the non-random allocation of students into classes of differing sizes, in order to estimate an education production function we need an exogenous factor that determines class size, but is orthogonal to unobserved determinants of educational outcomes. We use state special education policy to instrument for class size which draws upon the growing importance of special education to the variation in class size. Special education has become an increasingly prevalent program in schools over the past twenty years. In 1976, 8.3% of total enrollment was classified as "special needs"; the percentage had grown to 11% by 1988 (Digest of Education Statistics, 1991).[24] The largest growing component is those labelled "learning

[23] Indeed, if schools seek to maximize test scores (subject to cost constraints), then were these positive estimates in fact causal, given that smaller classes are more costly for schools, we should expect to see schools increase class sizes. The fact that we tend to see richer school districts set lower class sizes than poorer class sizes is, thus, prime facie evidence that the causal impact of class size is not positive, although it might well be zero.

[24] We have not found similar figures for remedial education. We suspect this is because states do not regulate "remedial education" as they do "special needs" leaving districts to design their own policies.

disabled" which grew from 1.8% to 4.9% over the same period. As noted in Table 2, the class sizes of special needs (and remedial classes) are roughly one-half those of other types of classes. This time-series trend could be responsible for the aggregate reduction in the average class size noted by researchers such as Flyer and Rosen (1994).

Actual class size will be correlated with the state regulations to the extent that schools base the entire structure of their class sizes on such state policy.[25] The relationship between the state maximum policies and actual 8th grade class size are illustrated in the regressions in Table 5 below:

Table 5
Regression of 8th Grade Class Size on State Special Education Maximum Class Size Policy

	Special Education Category	
	Emotionally Disturbed	Learning Disabled
Log Maximum Class Size	-4.566 (0.738)	-5.243 (1.545)
Log Maximum Class Size, Squared	1.197 (0.149)	1.270 (0.282)
No State Policy	-0.730 (0.107)	-0.830 (0.103)
F-Statistic for Instruments	84.96	81.25

Notes: There are 20,131 observations. See note to Table 6 for other regressors.

As the maximum state class size for special education classes increases, the student's class size initially decreases, but begins to increase after reaching a minimum of approximately two students. (Overall, the instruments are positively correlated with the actual class size.) Further, the F-statistics

[25] Our state policies come from Staff to Student Ratios: Class Size/Caseload (1986) and Staff to Student Ratios: Class Size/Caseload Supplement (1989) from the National Association of State Directors of Special Education, Inc. Washington, D.C.

suggest that these instruments are highly correlated with the endogenous variable (Bound, Jaeger, Baker (1993)). In our analysis we include the logarithm of the maximum class size, the logarithm of the maximum class size squared, and dummies indicating that the state does not regulate the maximum class size for five of the seven special education categories, generating 15 instruments.[26] All of the first-stage coefficients are presented in Appendix Table 2.

Using state regulations has the disadvantage of potentially reflecting the preferences of the residents (voters) of the state towards education rather than exogenously determined maximum class sizes. We control for this possibility, to some extent, by including four region dummies, state average income per capita, state expenditures per student, and the percentage of the state population twenty-five years and older with four or more years of college.[27] A second disadvantage is that the identifying information only varies at the state level. As the analysis reveals, the fact that we have on average 800 students per state may confound the interpretation of our diagnostic statistics.

Effects on Test Scores

Our basic results are presented in Table 6. This table shows the estimates from regressions of 10th and 12th grade test scores on a variety of explanatory variables, including the class size for the student's class for the subject of the test. Columns (2), (4), (6), and (7) also control for the student's previous test score(s) to account for previous school quality. The OLS results in Table 6 suggest that class size has a positive and significant effect on test scores. Increasing the class size by one student in the 8th grade increases the 10th grade test score 0.05 points above and beyond the student's 8th grade test score in the same subject (using the results from column 2), which is almost 1% of a standard deviation of the

[26] We use the logarithm, although we get similar results using a quadratic in the levels.

[27] Expenditures are for the 1987-1988 school year and are reported in the Digest of Education Statistics (1991), state educational attainment is from the Statistical Abstract (1993).

Table 6

OLS and IV Estimates of 10th and 12th Grade Test Scores
Using Actual 8th and 10th Grade Class Size

	Dependent Variable						
	10th Grade Test Score		12th Grade Test Score				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				OLS			
8th Grade Class Size	0.138 (0.012)	0.046 (0.007)			0.130 (0.017)	0.047 (0.012)	0.018 (0.009)
10th Grade Class Size			0.043 (0.015)	0.021 (0.008)	0.034 (0.015)	0.014 (0.010)	0.019 (0.008)
8th Grade Test Score		0.948 (0.005)				0.877 (0.008)	0.198 (0.010)
10th Grade Test Score				0.873 (0.005)			0.731 (0.009)
p-value of F-stat. for Jt. Sign. of Class Sizes					0.000	0.000	0.002
R^2	0.494	0.823	0.549	0.874	0.555	0.802	0.879
				IV			
8th Grade Class Size	-0.422 (0.086)	-0.137 (0.049)			0.105 (0.155)	-0.122 (0.092)	-0.054 (0.072)
10th Grade Class Size			-0.876 (0.192)	-0.292 (0.092)	-0.929 (0.217)	-0.259 (0.126)	-0.241 (0.099)
8th Grade Test Score		0.957 (0.005)				0.891 (0.009)	0.205 (0.011)
10th Grade Test Score				0.877 (0.006)			0.733 (0.010)
p-value of F-stat. for Jt. Sign. of Class Sizes					0.000	0.016	0.017
GMM χ^2	124.812	62.406	66.362	18.664	61.177	40.439	19.701
1st-stage F-statistic/ 8th Grade Class Size	27.52	28.58			13.40	13.37	13.50
1st-stage F-statistic/ 10th Grade Class Size			5.52	5.62	4.79	4.79	4.96
No. of Observations	20131		10369				

Notes: Standard errors are in parentheses. Cols. (1) and (2) are weighted using the first follow-up student panel weights; cols. (3)-(7) are weighted using the second follow-up student panel weights. Other regressors include: a constant, dummies indicating the subject, family income, family income squared, dummies indicating if the students' parents are married, whether there is a computer in the household, and the size, urbanicity and region of the junior and/or senior high school, the percentage of students from single parent households in the junior and/or senior high school and dummies indicating if these percentages are missing; state expenditures per student (1987/1988 school year), average income, and percentage of population (twenty-five years and older) with four or more years of college. Instruments are the logarithm of the state maximum class sizes for special education classes by category and the logarithm of the maximum class size squared; see text. The critical value for a χ^2 with 14 degrees of freedom (for cols. 1-4) is 23.69 (at the 5% level), and for 13 degrees of freedom (for cols. 5-7) is 22.36.

overall test score gains. The coefficients, however, change sign once we instrument for class size. In the first five columns the coefficients on class size are negative and significant. When both the 8th and 10th grade class sizes are included, the coefficient on the 10th grade class size is negative and statistically significant; and while the coefficient on the 8th grade class size is often insignificant, the coefficients on both class sizes are jointly significant.[28] Such results suggest that students are not randomly allocated to classes and that students in larger classes have lower test scores. Our results are qualitatively similar to those found by Blinder (forthcoming) who also uses the NELS and instruments for actual class size using the average class size in the school.[29]

We also present the χ^2 statistic for the generalized-method-of-moments (GMM) test of overidentification (Newey (1985)). The test involves regressing the residuals from the second-stage equation on the first-stage instruments. The R^2 of this regression times the number of observations is distributed as a χ^2 with k-q degrees of freedom where k is the number of instruments and q the number of endogenous variables. The null hypothesis is that the instruments give a consistent set of IV estimates. In all but two cases, we reject the null hypothesis. It is possible that because there are so many observations per state, even instruments that are only weakly correlated with the second-stage residual will generate an R^2 large enough to reject the null hypothesis. Unfortunately, the 800 observations per state

---

[28] Because 92% of the first follow-up and 55% of the second follow-up samples have two records, there is an individual component to the error term, and the stratified sampling results in a school component to the error term. We have accounted for neither of these effects in the reported standard errors. Our best guess (based on random- and fixed-effects) is that the effect of the individual component on our standard errors is negligible, but that the failure to account for the school effects may downward bias the reported standard errors by as much as 100%.

[29] Blinder presents results for test score levels rather than gains. However, the cross-school variation in class size may be contaminated by Tiebout sorting of "good" students into schools with smaller average class sizes. We argue that modeling test score gains (or including the lagged test score as an explanatory regressor) rather than levels helps to mitigate this source of bias.

may also have generated the relatively large F-statistics in the first-stage equations. Consequently, we interpret our results as suggestive, but not conclusive.[30]

We turn to the results of a "true" experiment of the effect of class sizes on test score outcomes in an effort to corroborate our IV estimates. A statewide experiment was conducted in Tennessee in which students in kindergarten and first grade were randomly assigned to small classes (average size of 15 students), regular classes (average size of 23 students), and regular with an aide (also an average size of 23 students, but with a teacher's aide). The results are reported in Finn and Achilles (1990). Finn and Achilles compare "small" classes with "average" classes (where "average" includes all classes of size 22-25 with and without a teacher's aide) when reporting their findings for the cross-sectional results for the end of Grade 1 reading and math test scores (Table 6 of their paper).[31] Finn and Achilles found that for minorities, a reduction of 8 students per class, on average, resulted in an increase of 0.35σ for reading (where σ is the standard deviation of the minority reading test score distribution) and 0.23σ for math. For

[30] We also exploited the panel structure of the NELS data and estimated a first-differenced version of equation (2) which nets out any time-invariant individual level heterogeneity. Specifically, we regressed the change in the test score on the change in the class size, and included other X's to control for observable heterogeneity in the changes. As usual, the standard errors from this type of differencing were greatly magnified relative to OLS on the levels (they increased by about a factor of 6). The coefficient magnitudes were positive, about 2 to 3 times larger than those from OLS, and quite insignificant. While the first-differenced specification is appealing in that it differences out the previous history of school quality and any time-invariant individual heterogeneity, it still suffers from the non-random assignment of students to classes. If 8th grade classes are not randomly assigned and 8th grade class size interacts with unobservables in affecting achievement, the first-difference will not solve the selection problem. Thus, we were not surprised to obtain results generally consistent with those from OLS.

[31] The longitudinal analysis at the end of their paper does not report the effect sizes in terms of the standard deviation of the gain scores, and so do not allow for direct comparison with our outcome in terms of its standard deviation. The longitudinal results lead, however, to the same qualitative conclusions as the cross sectional results.

white students, the effects were smaller, 0.13σ for reading and 0.15σ for math, but still statistically significant.[32]

We compare our results to the Tennessee experiment by estimating the implied effect of lowering the average class size by 8 students using our estimated coefficients. We obtain an implied increase in the average test score gain of (-0.2)*(-8) = 1.6 or approximately 0.29σ, where σ is the standard deviation of the entire test score gain distribution.[33] Thus, our findings from the IV strategy suggest an effect size that is comparable to those derived from a "true" experiment using actual class size.[34]

Effects on Drop Out Behavior

In Table 7 we examine the effect of class size on another academic outcome, the event of dropping out between 8th and 10th grade and between 8th and 12th grade. We follow the structure of reporting results used in Table 6, and again include a lagged test score as an attempt to capture the entire past history of schooling inputs. The OLS results reveal a statistically significant negative relationship between class size and subsequent drop out activity, suggesting that a reduction in class size would actually increase drop out activity. The IV results indicate that the causal effect of class size on dropout activity is essentially zero, although the fact that the standard error is so much larger than that for the OLS estimates leads us to interpret the insignificance of the point estimates in columns (2) and (4) with caution:

---

[32] The larger effect sizes for minorities may be partially due to the fact that these results do not condition on family background because the randomization was not within racial strata. If it had been, we might expect more similar effect sizes for the two races.

[33] The standard deviation is approximately 5.6; the -0.2 estimate is roughly the average point estimate of the effect of class size after controlling for the base year score.

[34] We can also compare our results to Card and Krueger's (1992) study of the effects of statewide differences in pupil-teacher ratios on differences in earnings of black and white workers educated in those states. They found that a statewide reduction in pupil-teacher ratios of 8 pupils would lead to an implied effect size of 0.16σ (using the results in their Table X) where σ is the standard deviation of the difference in mean log earnings for blacks and whites.

Table 7

Linear Probability and IV Estimates of the Likelihood of Dropping Out Using Actual 8th and 10th Grade Class Size

	Dependent Variable					
	Ever Dropped Out Between 8th and 10th Grade		Ever Dropped Out Between 8th and 12th Grade			
	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Avg. 8th Grade Class Size (÷10)	-0.025 (0.003)	-0.021 (0.004)			-0.013 (0.007)	-0.005 (0.007)
Avg. 10th Grade Class Size (÷10)			-0.023 (0.006)	-0.015 (0.006)	-0.020 (0.006)	-0.014 (0.006)
8th Grade Test Score (÷10)		-0.027 (0.003)		-0.078 (0.005)		-0.077 (0.005)
p-value of F-stat for Jt. Sign. of Class Sizes					0.000	0.035
R^2	0.035	0.043	0.031	0.059	0.041	0.067
			IV			
Avg. 8th Grade Class Size (÷10)	-0.006 (0.022)	-0.014 (0.021)			0.027 (0.044)	0.027 (0.043)
Avg. 10th Grade Class Size (÷10)			-0.009 (0.065)	-0.035 (0.063)	-0.017 (0.080)	-0.043 (0.077)
8th Grade Test Score (÷10)		-0.027 (0.003)		-0.077 (0.007)		-0.077 (0.008)
p-value of F-stat. for Jt. Sign. of Class Sizes					0.825	0.758
GMM χ^2	95.141	94.095	23.055	23.055	19.177	19.177
1st-stage F-statistic/ 8th Grade Class Size	19.56	20.89			14.23	14.51
1st-stage F-statistic/ 10th Grade Class Size			4.40	4.66	3.21	3.41
No. of Observations	10455		7685		7671	

Notes: Standard errors are in parentheses. Cols. (1) and (2) weighted using the first follow-up student panel weights; cols. (3)-(6) are weighted using the second follow-up student panel weights. Other regressors include: a constant, dummies indicating the subject, family income, family income squared, dummies indicating if the students' parents are married, whether there is a computer in the household, the size, urbanicity and region of the junior and/or senior high school, the percentage of students from single parent households in the junior and/or senior high school and dummies indicating if these percentages are missing; state expenditures per student (1987/1988 school year), average income, and percentage of population (twenty-five years and older) with four or more years of college. Instruments are the logarithm of the state maximum class sizes for special education classes by category and the logarithm of the maximum class size squared; see text. The critical value for a χ^2 with 14 degrees of freedom (for cols. 1-4) is 23.69 (at the 5% level), and for 13 degrees of freedom (for cols. 5-6) is 22.36. The (weighted) mean of the dependent variable in columns (1) and (2) is 3.06%; the (weighted) mean for columns (3)-(6) is 8.59%.

the causal effect may still be negative, even though our IV strategy cannot detect it. But by either set of results it is difficult to make the case that a larger class size leads causally to more dropout activity.

Since most students turn age-eligible to drop out in the 10th grade, the students who drop out between 8th and 10th grade are perhaps best thought of as constrained (by the legal age requirement) in their decision to drop out of school. It is therefore plausible that a class size reduction in 8th grade would not suffice to discourage a student from dropping out, given that the intervention occurs at such a late date.[35] Put differently, there may be treatment effect heterogeneity for class size interventions, and students at the bottom of the achievement distribution may not be affected much by their 8th grade class size. The fact that the OLS coefficients are quite significantly negative is probably due to the fact that the causal treatment effect is either zero or slightly negative, and so the OLS coefficients largely reflect the selection effect.[36] On the whole, we probably could not learn much about the effect of class size on dropouts, unless we were to have access to early school and class quality measures.

V. Implications for Black-white Achievement Differences

Our results indicate that one cannot use actual class size without accounting for the selection process into the classes of differing sizes and that once addressed, there is evidence of a negative effect of larger classes on test scores. Furthermore, blacks appear to be in larger classes conditional on class type. In this section we attempt to gauge the extent to which using actual class size could account for

[35] See Boozer (1992) for evidence that age-at-grade is a quite significant indicator of dropout status, and some further evidence from the High School and Beyond that this age measure is substantially determined by early schooling environments.

[36] Interestingly, Boozer (1992) found that a school's pupil-teacher ratio was in most cases positively associated with a student's greater propensity to drop out of school; a school's fraction black enrollment was also highly associated with dropping out activity. If school level averages are more highly correlated over time than are the individual class sizes, such a result may not be surprising since early schooling appears crucial for drop out behavior. In this case, school level averages would serve as a better proxy for early schooling than the individual class size.

differences in black and white educational achievement. In Table 8 we estimate separate IV regressions for black, white, and hispanic students and then estimate the test score gains of each using the average class size for blacks, whites, and hispanics, keeping the other variables at their original values. Estimating separate regressions by race reveals that the result in Table 6 for the 10^{th} grade test score is primarily driven by the effect of class size for black students. However, the results for the gain in scores between 10^{th} and 12^{th} grade suggest that white students experience more of an increase in test score gains for a given reduction in class size than do black students.

The IV results indicate that moving blacks into the average class size of white students in the 8^{th} grade (i.e., decreasing class size by about one student) would increase their test score gains in 10^{th} grade by about 6%, thereby closing the black-white test score gap (of gains) by about 15%. The effects for the gains from 10^{th} to 12^{th} grade are considerably smaller in that a one student decrease would lead to a 1% increase in gains for black students which is only 4% of the black-white gap.

VI. Conclusion

In our analysis we only address differences in class size and, as such, cannot speak to school quality more generally.[37] Nevertheless, our results indicate that some of our conventional wisdom regarding school quality may be misleading. We use measures of average class size and actual class size and find that black students are in larger classes than white students. We also find that while, ideally, researchers should use actual class size in education production functions, using actual class size in an OLS regression exacerbates non-random allocation of students to classes of differing sizes.[38] When this

---

[37] That said, 21% (a plurality) of teachers in the NJ Survey reported that hiring more teachers to reduce class size should be the highest priority were the school to receive additional funds. Buying more computers was a close second with 18.6%.

[38] On the other hand, actual size may also be subject to more classical measurement error than aggregate measures (Card and Krueger (1994)).

Table 8

Using Actual Class Size to Explain Racial Differences in Test Score Gains: IV Estimates Using State Special Education Class Size Policy as an Instrument for Class Size (Evidence from the NELS)

	Dependent Variable: 8th-10th Grade Test Score Gain			Dependent Variable: 10th-12th Grade Test Score Gain		
	Black	White	Hispanic	Black	White	Hispanic
8th Grade Class Size Coefficient (Std. Error)	-0.223 (0.102)	-0.005 (0.055)	0.261 (0.092)			
10th Grade Class Size Coefficient (Std. Error)				-0.071 (0.138)	-0.145 (0.093)	0.178 (0.138)
Mean Test Score Gain Using Black Mean Class Size	2.846 [2.479]	3.916 [2.260]	3.347 [2.217]	2.326 [1.468]	2.661 [1.016]	2.752 [1.101]
Mean Test Score Gain Using White Mean Class Size	3.010 [2.101]	3.920 [2.262]	3.137 [2.217]	2.341 [1.412]	2.684 [1.299]	2.718 [1.101]
Mean Test Score Gain Using Hispanic Mean Class Size	2.749 [2.101]	3.914 [2.260]	3.444 [2.621]	2.265 [1.412]	2.533 [1.016]	2.910 [1.587]
F-statistics for Instruments from First-Stage	5.02	22.79	8.33	2.26	5.14	2.69
GMM χ^2	21.719	55.757	21.013	18.640	18.708	24.768
Mean Class Size	25.067	24.263	25.431	23.591	23.402	24.473
No. of Observations	1922	14673	2284	951	7795	1032

Notes: Standard deviations in brackets, unless otherwise noted. The regressions and means in cols. (1)-(3) are weighted using the first follow-up student panel weights; cols. (4)-(6) use the second follow-up student panel weights. Other regressors include: a constant, the 8th grade or 10th grade test score, dummies indicating the subject, family income, family income squared, dummies indicating if the students' parents are married, whether there is a computer in the household, and the size, urbanicity and region of the junior or senior high school, the percentage of students from single parent households in the junior or senior high school and dummies indicating if these percentages are missing; state expenditures per student, average income, and percentage of population with four or more years of college. Instruments are the logarithm of the state maximum class sizes for special education classes by category and the logarithm of the maximum class size squared; see text. The critical value for the χ^2 with 14 degrees of freedom is 23.69 (at the 5% level). The comparable OLS regression coefficients (standard errors) for the 8th grade class size are 0.036 (0.019) for blacks, 0.042 (0.008) for whites, and 0.044 (0.021) for hispanics; and for the 10th grade class size are 0.004 (0.026) for blacks, 0.020 (0.009) for whites, and -0.002 (0.027) for hispanics.

source of selection is addressed, we find and report on evidence that smaller classes can increase educational achievement and can account for some of the black-white differences in educational achievement.

Unfortunately, our results cannot directly assess the impact of class size on labor market outcomes. However, Grogger (1994), using the High School and Beyond (HSB) and the National Longitudinal Study of the High School Class of 1972 (NLS-72), estimated a significant relationship between test score levels and wage outcomes. He found that math, vocabulary, and perception test scores accounted for 1/3 to almost 1/2 of the black-white wage gap. Although the age of our sample prohibits such a direct investigation as this, the similarity of the testing instrument in the NELS and the HSB data suggests that our results for test scores may well have labor market implications. Due to differences in specifications, however, we cannot gauge the potential magnitude of the impact.[39]

Finally, our results suggest that the growth of special education merits greater attention. For example Boozer, Krueger, Wolkon (1992) report a monotonic fall in the black-white difference in the pupil-teacher ratio since 1915 suggesting that the variation in pupil-teacher ratio may have decreased since the first half of this century. Flyer and Rosen (1994) note that only about one-half of the decrease in the pupil-teacher ratio from 1961-1991 can be accounted for by average class size, and that the discrepancies are "impossible" to resolve (p. 37). However, smaller special education classes (and compensatory education, in general) combined with larger "regular" classes may explain such trends, especially if the growth of special education has been unevenly distributed across schools.

---

[39] Grogger's specifications are the log wage on the level of the test scores, whereas we work with the impact of class size on the test score gain. In addition, the test subject areas only overlap in mathematics. Thus, while the dramatic effect of the inclusion of test scores on the black-white wage gap (documented more recently by Neal and Johnson (1994) using the AFQT scores in the NLSY) is highly suggestive that the effects we find of class size on test scores would translate to effects of class size on earnings, we cannot quantitatively assess the magnitude of such an effect.

References

Blinder, Karen. "Does Class Size Matter?" Economics of Education Review, forthcoming.

Boozer, Michael. "Segregation, Dropouts, and School Quality," mimeo, Princeton University Department of Economics, 1992.

Boozer, Michael A., Alan B. Krueger and Shari Wolkon. "Race and School Quality Since Brown v. Board of Education." Brookings Papers: Microeconomics 1992, 269-338.

Card, David and Alan B. Krueger. "School Quality and Black-White Relative Earnings: A Direct Assessment," Quarterly Journal of Economics, 107 (February, 1992).

Card, David and Alan B. Krueger. "The Economic Return to School Quality: A Partial Survey," Working Paper #334, Industrial Relations Section, Princeton University, 1994.

Digest of Education Statistics, Washington, D.C.: U.S. Department of Health, Education, and Welfare, Education Division, National Center for Education Statistics, 1991.

Finn, Jeremy D., and Charles M. Achilles. "Answers and Questions About Class Size: A Statewide Experiment," American Educational Research Journal, 27 no. 3 (Fall, 1990): 557-577.

Flyer, Frederick, and Rosen, Sherwin. "The New Economics of Teachers and Education," NBER Working Paper # 4828.

Grogger, Jeff. "Does School Quality Explain the Recent Black/White Wage Trend?" mimeo, UC Santa Barbara Department of Economics, 1994.

Hanushek, Eric A. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," Journal of Human Resources, 14 (Summer, 1979): 351-388.

Hanushek, Eric A. and John F. Kain. "On the Value of Equality of Educational Opportunity as a Guide to Public Policy," in On Equality of Educational Opportunity, Frederick Mosteller and Daniel P. Moynihan, editors (New York: Random House, 1972).

Juhn, Chinhui, Kevin Murphy and Brooks Pierce. "Accounting for the Slowdown in Black-White Wage Convergence." In Marvin H. Kosters, (ed.) Workers and their Wages: Changing Patterns in the United States (Washington, D.C.: AEI Press, 1991).

Kozol, Jonathan. Savage Inequalities: Children in America's Schools (New York: Crown Publishers, 1991).

National Education Longitudinal Study of 1988, Second Follow-Up: School Component Data File User's Manual, Washington D.C.: National Center for Education Statistics, 1994.

Neal, Derek A. and William R. Johnson. "The Role of Pre-Market Factors in Black-White Wage Differences," mimeo, University of Chicago Department of Economics, 1994.

Newey, Whitney. "Generalized Method of Moments Specification Testing," Journal of Econometrics, 29 (1985): 229-256.

Smith, James P. and Finis Welch. "Black Economic Progress After Myrdal," Journal of Economic Literature 27 (June, 1989): 519-564.

Staff to Student Ratios: Class Size/Caseload, Washington D.C.: National Association of State Directors of Special Education, Inc., 1986.

Staff to Student Ratios: Class Size/Caseload Supplement, Washington D.C.: National Association of State Directors of Special Education, Inc., August 1989.

U.S. Bureau of the Census, Statistical Abstract of the United States: 1993 (113th edition.) Washington, D.C., 1993.

Data Appendix
Table 1

Mean Teacher Characteristics:
All Teachers in New Jersey vs. New Jersey Survey

	Data	
	All Teachers in NJ[a]	NJ Survey
Highest Degree		
BA	0.629 (0.002)	0.542 (0.024)
MA	0.356 (0.002)	0.440 (0.024)
PhD	0.007 (0.0003)	0.009 (0.005)
Years of Teaching Experience		
Total	15.480 (0.032)	16.572 (0.433)
In the District	13.112 (0.033)	12.812 (0.427)
Type of Teacher		
Special Education	0.129 (0.001)	0.145 (0.017)
Remedial Education	0.053 (0.001)	0.027 (0.006)
No. of Observations	77,227	441

Notes: Standard errors are in parentheses.

[a] Source: The New Jersey Department of Education

Data Appendix
Table 2

Mean Characteristics of Schools in New Jersey Sample Compared to All Schools in New Jersey

	All Schools in NJ		Schools in NJ Survey[a]	
	Mean	Std. Error	Mean	Std. Error
Pupil-Teacher Ratio (Common Core)	15.58	0.105	16.43	0.193
Pupil-Teacher Ratio (NJ Dept of Educ)	15.21	0.092	15.86	0.174
Total Number of Students	488.59	7.098	425.31	13.705
% White	0.68	0.007	0.74	0.014
% Black	0.16	0.005	0.11	0.010
% Hispanic	0.10	0.004	0.08	0.009
# Classroom Teachers	31.47	0.530	28.15	1.028
# Staff	39.05	0.665	32.80	1.237
Avg. Salary	42263.91	118.97	42441.73	306.41
Avg. Yrs. Experience (Total)	15.29	0.071	15.45	0.173
Avg. Yrs. Experience (NJ)	14.69	0.071	14.88	0.174
Avg. Yrs. Experience (District)	12.88	0.074	13.21	0.177
% Administrators	0.06	0.001	0.04	0.001
% Other Administrators	0.03	0.001	0.01	0.001
% Remedial Teachers	0.04	0.001	0.03	0.002
% Bilingual Educ Teachers	0.02	0.001	0.01	0.002
% Academic Teachers	0.20	0.004	0.16	0.010
% Vocational Teachers	0.03	0.001	0.01	0.002
% Other Teachers	0.06	0.001	0.05	0.002
% Special Ed	0.11	0.002	0.11	0.004
% Supp. Services	0.12	0.002	0.08	0.003
% General Elementary	0.40	0.006	0.54	0.014
Elementary School	0.77	0.009	0.88	0.017
Jr. High School	0.29	0.009	0.25	0.023
High School	0.15	0.007	0.09	0.016
Urban	0.20	0.008	0.13	0.018
Current Expenditures/Pupil	8937.73	33.80	8905.30	103.33
Total Expenditure/Pupil	9735.48	36.05	9690.54	108.19
No. of Observations	2550		349	

Notes: The number of observations varies according to data availability.
[a] Weighted by the inverse of the number of teachers in the school.

Data Appendix:
Assessing the Reliability of Teachers' Responses in the NJ Survey

We are able to assess the accuracy of the teachers' responses, particularly with regard to school level characteristics, such as the total number of teachers and total enrollment because in 76 of the 389 useable schools, more than one teacher was surveyed.[40] Thus, we are able to compute reliability ratios for several variables using these multiple measures. Let x_1 and x_2 be two reports of the true value x and let $x_1 = x + \varepsilon_1$ and $x_2 = x + \varepsilon_2$ where ε_1 and ε_2 are the measurement errors in the reports. We assume that ε_1 and ε_2 are uncorrelated with the x and with each other, and that $var(x_1)=var(x_2)$. Under this model, we calculate the reliability ratio, λ, as the correlation between the two reports of the variables so that $\lambda = cov(x_1,x_2)/[var(x_1)\, var(x_2)]^{1/2} = var(x)/var(x_1)$, or the proportion of the variance in the reported measure that is accounted for by the true variation in x. The results are presented in Table 3, below:

Table 3
Reliability Ratios for Measures in the New Jersey Survey

Variable	Reliability Ratio
School Avg. Class Size	0.787
Number of (FT) Teachers	0.722
Total School Enrollment	0.852
Pupil-Teacher Ratio	0.012
Percentage (School) Black	0.951
Percentage (School) Hispanic	0.918

These estimates suggest that about 20% of the measured variance in average school class size, 28% of the measured variance in total number of teachers, and 15% of the measured variance in total enrollment is error. The reliability ratios for the racial composition of the school are much higher, 5-8% of the observed variance being attributable to mis-measurement. On the other hand, when the total school and enrollment and the number of teachers are combined to generate the pupil-teacher ratio, fully 99% of the observed variance is due to error. The large change in the reliability of the ratio of the two measures is due to the fact that the measurement error in the pupil-teacher ratio is a function of the measurement errors in the number of teachers and enrollment weighted by the

[40] There were 58 schools in which two teachers were surveyed and 18 schools in which three teachers were surveyed.

variances of these two measures.[41] Small amounts of measurement error in the individual components of the pupil-teacher ratio can have devastating effects on the ratio. Thus, while the teachers appear to have reasonably accurate perceptions about the school characteristics, the pupil-teacher ratio in the NJ Survey may be severely mismeasured.

---

[41] The reliability ratio for the pupil-teacher ratio is:

$$\lambda_{PT} = \frac{a\lambda_P var(P_1) - 2bcov(P,T) + c\lambda_T var(T_1)}{a_1 var(P_1) - 2b_1 cov(P_1,T_1) + c_1 var(T_1)}$$

λ_P is the reliability ratio for total enrollment, λ_T is the reliability ratio for the number of teachers, P_1 and T_1 are observed measures of total enrollment and the number of teachers respectively, and P and T are the true values; $a=(1/(E(T))^2)$, $b=(E(P)/(E(T))^3)$, and $c=((E(P))^2/(E(T))^4)$ (using the true values), and a_1, b_1, and c_1 are similar constants using the observed values.

Appendix Table 1
NELS Sample Descriptive Statistics (Weighted)

	1st FU Sample		2nd FU Sample	
	Mean	Std. Dev.	Mean	Std. Dev.
8th Grade Class Size	24.506	5.672	24.438	5.611
10th Grade Class Size			23.542	6.581
8th Grade Test Score	27.731	9.952	28.774	10.584
10th Grade Test Score	31.501	12.131	33.104	12.785
12th Grade Test Score			35.788	13.588
Black	0.106	0.308	0.100	0.300
Hispanic	0.088	0.283	0.077	0.266
Female	0.496	0.500	0.498	0.500
Reading	0.252	0.434	0.278	0.448
History	0.248	0.432	0.183	0.387
Math	0.252	0.434	0.281	0.450
Family Income	36793.000	30814.350	38087.140	30262.420
Family Income Squared (÷ 1000000)	23003.202	5799.687	2366.356	5745.716
Has a Home Computer	0.424	0.494	0.440	0.496
Parents are Married	0.801	0.399	0.833	0.373
% 8th Grade School with Single Parent	27.594	17.731	27.222	17.728
% 8th Grade School with Single Parent, Don't Know	0.044	0.205	0.046	0.210
% 8th Grade School with Single Parent, Missing	0.005	0.068	0.002	0.043
8th Grade School Urban	0.183	0.386	0.151	0.358
8th Grade School Rural	0.379	0.485	0.406	0.491
8th Grade School Total Enrollment	686.073	347.904	675.110	328.302
8th Grade School in North Central	0.280	0.449	0.283	0.450
8th Grade School in North East	0.174	0.379	0.187	0.390
8th Grade School in South	0.373	0.484	0.377	0.485
% 10th Grade School with Single Parent			23.898	17.009
% 10th Grade School with Single Parent, Missing			0.143	0.350
10th Grade School Urban			0.158	0.364
10th Grade School Rural			0.405	0.491
10th Grade School Total Enrollment			1152.548	642.713
10th Grade School in North Central			0.282	0.450
10th Grade School in North East			0.188	0.390
10th Grade School in South			0.377	0.485

Appendix Table 1 (cont.)
NELS Sample Descriptive Statistics (Weighted)

	1st FU Sample		2nd FU Sample	
	Mean	Std. Dev.	Mean	Std. Dev.
State Expenditures/Student	4393.784	1127.230	4400.634	1092.951
State Income/Capita	12.776	2.040	12.709	2.015
State Educational Attainment	15.691	2.648	15.565	2.681
Emotionally Disturbed, Max.	2.286	0.308	2.280	0.300
Emotionally Disturbed, Max. Squared	5.322	1.540	5.288	1.450
Emotionally Disturbed, No Policy	0.310	0.462	0.278	0.448
Learning Disabled, Max.	2.477	0.261	2.470	0.250
Learning Disabled, Max. Squared	6.202	1.427	6.163	1.344
Learning Disabled, No Policy	0.338	0.473	0.310	0.463
Hearning Impaired, Max.	2.286	0.325	2.266	0.314
Hearing Impaired, Max. Squared	5.333	1.603	5.234	1.503
Hearing Impaired, No Policy	0.359	0.480	0.327	0.469
Mentally Retarded, Max.	2.475	0.255	2.466	0.246
Mentally Retarded, Max. Squared	6.189	1.387	6.141	1.302
Mentally Retarded, No Policy	0.371	0.483	0.339	0.473
Visually Impaired, Max.	2.278	0.270	2.272	0.256
Visually Impaired, Max. Squared	5.262	1.419	5.227	1.305
Visually Impaired, No Policy	0.389	0.487	0.356	0.479
Number of Observations	20131		10369	

Appendix Table 2
First-Stage Regressions of Class Size on the State Maximum Class Size for Special Education

	Dependent Variable	
Instruments (Log State Max. Class Size)	8th Grade Class Size	10th Grade Class Size
Emotionally Disturbed	-27.229	-30.721
	(3.114)	(4.381)
Emotionally Disturbed, Squared	6.623	7.644
	(0.740)	(1.036)
Emotionally Disturbed, No State Policy	0.644	0.941
	(0.366)	(0.500)
Learning Disabled	-23.130	-36.006
	(5.472)	(7.233)
Learning Disabled, Squared	-4.922	-7.930
	(1.124)	(1.489)
Learning Disabled, No State Policy	-2.668	-2.591
	(0.327)	(0.452)
Hearing Impaired	1.823	1.756
	(1.313)	(1.775)
Hearing Impaired, Squared	-0.465	-0.628
	(0.318)	(0.430)
Hearing Impaired, No State Policy	-0.333	-0.951
	(0.263)	(0.364)
Mentally Retarded	-3.786	-14.379
	(4.372)	(5.761)
Mentally Retarded, Squared	0.783	3.038
	(0.924)	(1.218)
Mentally Retarded, No State Policy	1.702	1.595
	(0.332)	(0.448)
Visually Impaired	17.404	23.729
	(5.017)	(7.097)
Visually Impaired, Squared	-3.616	-4.702
	(1.096)	(1.540)
Visually Impaired, No State Policy	-0.408	-0.012
	(0.275)	(0.362)
F-Statistic	27.52	13.42
No. of Observations	20131	10369

Notes: Standard errors are in parentheses. Col. (1) is weighted using the first follow-up student panel weights; col. (2) is weighted using the second follow-up student panel weights. Other regressors include: a constant, dummies indicating the subject, family income, family income squared, dummies indicating if the students' parents are married, whether there is a computer in the household, and the size, urbanicity and region of the junior or senior high school, the percentage of students from single parent households in the junior or senior high school and dummies indicating if these percentages are missing; state expenditures per student (1987/1988 school year), average income, and percentage of population (twenty-five years and older) with four or more years of college.

Appendix Table 3

OLS and IV Estimates of 10th and 12th Grade Test Scores Using Actual 8th Grade Class Size and Second Follow-up Sample

	Does not include base year test score		Includes base year test score	
	OLS	IV	OLS	IV
8th Grade Class Size	0.134 (0.017)	-0.010 (0.121)	0.039 (0.010)	-0.143 (0.072)
8th Grade Test Score			0.929 (0.007)	0.938 (0.007)
R^2	0.523		0.838	
F-test of Excluded Variables from 1st Stage		13.42		13.35
GMM χ^2		103.690		40.439
No. of Observations	10369			

Notes: Standard errors are in parentheses. Weighted using the second follow-up student panel weights. Other regressors include: a constant, dummies indicating the subject, family income, family income squared, dummies indicating if the students' parents are married, whether there is a computer in the household, and the size, urbanicity and region of the junior high school, the percentage of students from single parent households in the junior high school and dummies indicating if this percentages is missing; state expenditures per student (1987/1988 school year), average income, and percentage of population (twenty-five years and older) with four or more years of college. Instruments are the logarithm of the state maximum class sizes for special education classes by category and the logarithm of the maximum class size squared; see text. The critical value for a χ^2 with 14 degrees of freedom (for cols. 1-4) is 23.69 (at the 5% level).

LISTED BELOW IS A SUMMARY OF RECENTLY PUBLISHED ECONOMIC GROWTH CENTER DISCUSSION PAPERS. COPIES ARE AVAILABLE AT $2.00 EACH PLUS POSTAGE BY WRITING TO THE PUBLICATIONS OFFICE, ECONOMIC GROWTH CENTER, P.O. BOX 208269, NEW HAVEN, CONNECTICUT 06520-8269.

699.	"Cooperatives as Information Machines: The Lending Practices of German Agricultural Credit Cooperatives, 1883-1914," August 1993. (35 pp.)	Timothy W. Guinnane
700.	"Regional Spillovers and Economic Growth," September 1993. (42 pp.)	Hak B. Chua
701.	"Understanding the Workhouse Test: Information and Poor Relief in Nineteenth-Century England," September 1993. (28 pp.)	Timothy Besley Stephen Coate Timothy W. Guinnane
702.	"Investments in the Schooling and Health of Women and Men: Quantities and Returns," August 1993. (70 pp.)	T. Paul Schultz
703.	"Marital Status and Fertility in the United States: Welfare and Labor Market Effects," September 1993. (39 pp.)	T. Paul Schultz
704.	"Regional Instability and Economic Growth: Thy Neighbor's Curse," October 1993. (38 pp.)	Hak B. Chua Alberto Ades
705.	"Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test," September 1993. (38 pp.)	Timothy Guinnane Abihijit Banerjee Timothy Besley
706.	"Indian Public Finance in the 1990s: Challenges and Prospects," December 1993. (58 pp.)	Willem H. Buiter Urjit R. Patel
707.	"Human Capital, Family Planning and Their Effects on Population Growth," January 1994. (12 pp.)	T. Paul Schultz
708.	"Measurement of the Total Factor Productivity of Spanish Agriculture: 1962-1989," December 1993. (15 pp.)	Robert E. Evenson M. Carmen Fernandez A. Casimiro Herruzo
709.	"Technological Change and Technology Strategy," January 1994. (132 pp.)	Robert E. Evenson Larry E. Westphal
710.	"Price Normalization and Equilibria in General Equilibrium Models of International Trade Under Imperfect Competition," April 1994. (19 pp.)	T.N. Srinivasan Kenneth Kletzer
711.	"Human Capital and Economic Development," May 1994. (31 pp.)	T. Paul Schultz
712.	"The World Bank Near the Turn of the Century," October 1994. (30 pp.)	Gustav Ranis

713.	"Genetic Resources, International Organizations, and Rice Varietal Improvement," July 1994. (38 pp.)	Robert E. Evenson Douglas Gollin
714.	"Productivity, Competitiveness, and Export Growth in a Less Developed Economy: A Study of Indian Punjab," August 1994. (27 pp.)	Lakhwinder Singh
715.	"Quality Improvements in Models of Growth," October 1994. (43 pp.)	Xavier Sala-i-Martin Robert J. Barro
716.	"Regional Cohesion: Evidence and Theories of Regional Growth and Convergence," October 1994. (46 pp.)	Xavier Sala-i-Martin
717.	"Lobbying and International Cooperation in Tariff Setting," June 1994. (28 pp.)	Philip I. Levy
718.	"A Political-Economic Analysis of Free Trade Agreements," June 1994. (36 pp.)	Philip I. Levy
719.	"Integrated Approaches to Human Resource Development," January 1995. (33 pp.)	T. Paul Schultz
720.	"Evaluation of Integrated Human Resource Programs," February 1995. (27 pp.)	T. Paul Schultz
721.	"Trade and the Environment: Does Environmental Diversity Detract from the Case for Free Trade?" January 1995. (100 pp.)	T. N. Srinivasan Jagdish Bhagwati
722.	"A Labor-Income-Based Measure of the Value of Human Capital: An Application to the States of the United States," December 1994. (61 pp.)	Xavier Sala-i-Martin Casey B. Mulligan
723.	"Measuring Aggregate Human Capital," January 1995. (48 pp.)	Xavier Sala-i-Martin Casey B. Mulligan
724.	"V-Goods and the Role of the Urban Informal Sector in Development," December 1994. (46 pp.)	Gustav Ranis Frances Stewart
725.	"World Income Components: Measuring and Exploiting International Risk Sharing Opportunities," March 1995. (69 pp.)	Robert J. Shiller Stefano Athanasoulis
726.	"The Limits to Land Reform: The Land Acts in Ireland, 1870-1909," June 1995. (31 pp.)	Timothy W. Guinnane Ronald I. Miller
727.	"Information and the Operation of Markets: Tests Based on a General Equilibrium Model of Land Leasing in India," June 1995. (37 pp.)	Jean Olson Lanjouw
728.	"Intraschool Variation in Class Size: Patterns and Implications," May 1995. (43 pp.)	Michael Boozer Cecilia Rouse