Spring April, 2009

Student Demographics, Teacher Sorting, and Teacher Quality: Evidence From the End of School Desegregation

Clement (Kirabo) Jackson, *Cornell University*

Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation

C. Kirabo Jackson, *Cornell University*

The reshuffling of students due to the end of student busing in Charlotte-Mecklenburg provides a unique opportunity to investigate the relationship between changes in student attributes and changes in teacher quality that are not confounded with changes in school or neighborhood characteristics. Comparisons of ordinary least squares and instrumental variable results suggest that spatial correlation between teachers' residences, students' residences, and schools could lead to spurious correlation between student attributes and teacher characteristics. Schools that experienced a repatriation of black students experienced a decrease in various measures of teacher quality. I provide evidence that this was primarily due to changes in labor supply.

I. Motivation and Introduction

Many education policy interventions, such as school vouchers, school choice, district consolidation, and student busing, change the student demographics of schools. Such policies are predicated, in part, on the hypothesis that it is helpful to reshuffle peers while keeping other things

I would like to thank Elias Bruegmann, Roland Fryer, Li Han, Caroline Hoxby, Clement Jackson, Dale Jorgenson, Lawrence Katz, Gauri Kartini Shastry, Katharine Emans Sims, Erin Strumpf, Daniel Tortorice, and participants at the National Bureau of Economic Research education group meeting for helpful suggestions and comments. I also thank Kara Bonneau of the North Carolina Education Research Data Center. All errors are my own. Contact the author at ckj5@cornell.edu.

[*Journal of Labor Economics,* 2009, vol. 27, no. 2]

0734-306X/2009/2702-0002$10.00

roughly the same. While this may be true, it may be impossible to keep teaching "roughly the same" if teacher quality is endogenous to student characteristics. Since salaries do not vary across schools within a district, teachers have little financial incentive to teach at undesirable schools. Since observably better teachers will be hired over weaker teachers and all teachers are likely to apply for the most desirable jobs, schools with undesirable working environments will have teachers of lower average quality. As such, if teachers prefer working environments with students of a particular demographic, teacher quality would be endogenous to student demographics and, ceteris paribus, students whom teachers find undesirable will be exposed to teachers of lower quality. With such teacher sorting, policies that change the composition of students may change the composition of teachers in unforeseen and undesirable ways. For example, the movement of high-quality teachers out of schools with growing black enrollment shares may be partially responsible for the ill effects of school segregation to black students documented by Guryan (2004) and Lutz (2005) and for the finding that higher black enrollment shares are associated with lower test scores (Hoxby 2000; Hanushek, Kain, and Rivkin 2004a).

Although research results are mixed, there is evidence that years of teaching experience, selectivity of undergraduate institutions, teachers' test scores, and regular licensure are associated with higher student achievement (Brewer and Ehrenberg 1994; Hanushek 1997; Brewer and Goldhaber 2000; Clotfelter, Ladd, and Vigdor 2006, 2007; Anthony and Goldhaber 2007). Studies that identify teachers associated with student test score gains show that a one standard deviation increase in teacher quality leads to between one-tenth and one-fifth of a standard deviation increase in math and reading scores (Rockoff 2004; Hanushek, Kain, and Rivkin 2005; Aaronson, Barrow, and Sander 2007). Jacob and Lefgren (2008) find that principals' subjective evaluations of teachers are highly correlated with subsequent increases in student achievement.

Researchers have found that high-poverty schools tend to have teachers with lower qualifications than low-poverty schools and that teachers tend to leave schools with low-achieving, poor, minority students, particularly when there are vacancies at schools with higher-achieving, affluent students. This evidence is based on observation of teacher attributes, or changes in teacher attributes, at schools whose student populations are unchanging or are changing because of unobserved factors that could also affect teacher labor supply. I provide an overview of the literature and discuss why, on the basis of previous studies, one cannot say whether the observed differences are caused by (*a*) school attributes that are correlated with student characteristics, (*b*) neighborhood attributes that are correlated with student characteristics, or (*c*) mobility of teachers toward their residences that happens to move them out of inner-city schools. Since previous studies have been unable to separate the effect of student characteristics on teacher quality from those of neighborhoods or schools, we have little knowledge of the direct relationship between student characteristics and teacher quality and little understanding of how policies that change the composition of students across schools might affect the distribution of teachers.

In 2002, the Charlotte-Mecklenburg (CM), North Carolina, school district ended its long-standing school integration policy, which entailed busing students across neighborhoods to maintain racial balance of the student bodies across schools. Since during busing CM schools were compelled to have student populations that were similar to the district average, between 2002 and 2003[1] the demographic makeup of schools quickly converged to that of their surrounding neighborhoods whereas other school attributes and neighborhood characteristics were largely unchanged.[2] The sudden changes in student attributes within schools over time that were due to the policy change provide a unique opportunity to observe teachers' reactions to exogenous changes in student attributes that were not correlated with changes in neighborhood or school characteristics. I exploit this policy change in order to overcome the limitations of previous studies to credibly uncover the causal effect of changes in student characteristics on changes in teacher quality.[3] Although a faculty desegregation order was issued in 1972, it had not been exercised for over 20 years, and there was no change in the district's hiring or teacher/principal placement practices over the sample period.[4] Also, CM has a policy of not forcibly relocating teachers across schools. As such, I interpret changes in teacher mobility to be primarily a labor supply response. Empirical evidence and conversations with district officials suggest that the changes in teachers' locations were not driven by changes in teacher demand; therefore, this analysis may provide empirical evidence of the sorting suggested by the theory of compensating differences.

[1] Throughout this paper I refer to school years by their ending calendar year. For example, the academic school year 2002–3 is referred to as 2003.

[2] In 2000–2001, only 48% of students in the county attended a school that deviated from the district average percentage of minority students by more than 15 percentage points. However, in 2004–5, after the policy change, that number increased to 74%. Source: NAACP, "Impact of Race Neutral Alternatives," http://www.naacpldf.org/ (accessed January 2007).

[3] Other researchers have used this policy change in CM as a way to study the effects of school choice on student outcomes (Hastings, Kane, and Staiger 2006; Hastings, Van Weelden, and Weinstein 2007; Hastings and Weinstein 2007) and to study the relationship between school characteristics and housing prices (Kane, Staiger, and Riegg 2005).

[4] Employees from the CM legal office, personnel office, and superintendent's office have all corroborated this statement. However, the superintendent did forcibly relocate two school principals after the sample period.

Since a racially integrated school in a predominantly black neighborhood would have experienced a larger repatriation of black students after busing ended than a predominantly black school in an identical neighborhood, I use the difference between the proportion of black students at the school and the proportion of black residents in its surrounding neighborhood *before the policy change* to predict the exogenous inflow/outflow of black students due to the policy change. While the policy change allows me to observe exogenous movement of students, I am unable to disentangle race from other student characteristics associated with race such as income level and achievement. Therefore, as in other studies, student race is a summary statistic for a variety of student attributes, and the results should be interpreted in that light.

I find that schools that had an inflow of black students as a result of the policy change had a decrease in the share of high-quality teachers, as measured by years of experience and certification test scores. Similarly to Hanushek et al. (2005), I use student achievement gains to estimate teacher value added, which I use as a measure of unobserved teacher quality. I find that schools that had an inflow of black students also experienced a decrease in average estimated teacher effectiveness in math and reading. These changes were largely driven by changes in the attributes of teachers who remained in these schools—indicating that experienced, high-scoring, and high-value-added teachers were relatively more likely to leave these schools. I find that black teachers were more likely to stay in these schools whereas white teachers were relatively unaffected, so that the percentage of black teachers increased. However, inflows of black students are associated with decreases in the average quality of both black and white teachers, suggesting that sorting by student race occurs both across and within teacher races. The relationship between teacher characteristics and student race differs in the within-school instrumental variables regressions and in the cross section, suggesting that some of the well-documented correlations are artifacts of residential segregation. This paper presents the first compelling evidence that the relationship between student demographics and teacher quality may be causal.

The data show that teachers in all CM schools were more likely to leave their current school and, more specifically, were more likely to move to other schools in CM the year before students were reassigned. Furthermore, the direction of the flow of black students is not correlated with the hiring of more teachers (vacancies). Both of these patterns suggest the changes were not demand driven and were instead due to a labor supply response. These patterns are consistent with a compensating differentials equilibrium in which teachers have heterogeneous tastes for student attributes so that teachers re-sorted in the face of an anticipated change in working conditions. The findings suggest that the widening black-white achievement gap associated with residential and school segregation and the negative relationship between student achievement and percentage of black students at the school are due, in part, to the endogeneity of teacher quality with respect to student characteristics. The findings underscore that policy makers should be careful to consider how teachers may reallocate themselves when students are moved across schools through vouchers, school choice, district consolidation, or student busing.

The remainder of the paper is structured as follows: Section II reviews the literature on teacher quality and student attributes. Section III describes the policy change and documents its effect on student characteristics. Section IV shows the effect of the policy change on teacher characteristics. Section V presents a graphical analysis of teacher turnover. Section VI uses disaggregated teacher data to explain the observed results in the aggregate, and Section VII presents concluding remarks.

II. Research on Student Attributes and Teacher Mobility

It has been well documented that inner-city, high-poverty schools with high ethnic minority enrollment shares tend to have teachers with lower qualifications than low-poverty schools (Betts, Rueben, and Danenberg 2000; Lankford, Loeb, and Wyckoff 2002; Hanushek, Kain, and Rivkin 2004b; Hanushek and Rivkin 2006; Clotfelter et al. 2007; Scafidi, Sjoquist, and Stinebrickner 2007). These researchers also found that low-income, inner-city schools experience higher teacher turnover, particularly among white teachers, than affluent, high-achieving suburban schools. While greatly informative, these studies compare the stock or the flow of teachers across schools in which student attributes are either unchanging or changing for reasons that may exert an independent effect on teacher labor supply decisions.[5]

Tracking the movement of teachers across schools, researchers have found that teachers, particularly those with more experience, in schools with low-achieving students move to higher-achieving schools—leaving districts that have high shares of low-income ethnic minority students with vacancies and unqualified instructors (Bohrnstedt and Stecher 1999; Lankford 1999; Betts et al. 2000; Lankford et al. 2002; Hanushek, Kain, and Rivkin 2004b; Hanushek et al. 2005). Hanushek, Kain, and Rivkin find that this movement is stronger for white teachers than for black teachers, suggesting that teachers may prefer own-race students.

[5] As noted by several researchers, attempting to separate the contribution of student attributes from those of school or neighborhood attributes (which are highly collinear and jointly determined) is a dubious exercise without independent exogenous variation. Although including school and neighborhood proxies can mitigate this problem, the strong collinearities among student demographics, school attributes, and neighborhood attributes render this solution unsatisfactory.

Analyzing New York teachers, Boyd et al. (2005) find that the geographic location of a school vis-à-vis where a teacher grew up plays a strong role in labor supply decisions. They find that teacher labor markets tend to be geographically small and that teachers express preferences to teach close to where they grew up, which in turn tends to be close to their current residences. The implications of the geospatial nature of teacher labor markets are that the spatial correlation between teachers' residential locations and those of the schools could generate both the cross-sectional relationship and the dynamics documented by researchers even if teachers have no preference for student or school attributes per se.

Consider the observations that experienced teachers leave inner-city schools when there are vacancies at affluent, suburban schools and that experienced teachers are less likely to teach at inner-city schools serving poor, minority populations. Since more experienced teachers are often given preference for new teaching positions, they have greater ability to express their preferences for schools. Since teachers—especially older teachers who are likely to have families—tend to live in suburban areas with reasonably good schools, their moving toward schools that are close to their homes will systematically move them out of inner-city schools that serve low-income, ethnic minority neighborhoods. In such a scenario, teachers' endogenous movements, especially those of experienced teachers, would be due to the spatial correlation between school demographics, neighborhood characteristics, and teachers' residential locations rather than a reflection of teachers' preferences for teaching at the schools per se. If the documented relationship between student and teacher attributes is an artifact of residential segregation, the interpretation of the evidence would be very different, as would policy prescriptions with regard to teacher recruiting and retention.[6]

To address this spatial correlation bias, one would like to observe changes in teacher labor supply decisions at schools in which student demographics are changing but for which the geospatial relationship between schools and their homes is unchanged. Given the limitations associated with observing endogenous movement of teachers across schools whose student populations are associated with a variety of other factors (including distance to home), the relatively sudden change in schools' student demographics caused by the end of the desegregation order in

[6] For example, policies that improve the quality of neighborhoods surrounding a school may make it easier to attract teachers to schools with large ethnic minority shares. Alternatively, policies that make it easier to live farther away from schools that are in undesirable neighborhoods could improve teacher retention. Schools could also actively recruit teachers who grew up close by or in similar neighborhoods. However, if teachers react to the demographics of students rather than to the neighborhoods of their schools, such policies would be largely ineffective.

CM may provide some new insights into the relationship between student attributes and teacher sorting.

III. The Policy Change and Its Effect on Student Characteristics

In 1971 the U.S. Supreme Court held that busing was an appropriate way to ensure that all students would receive equal educational opportunities regardless of their race (*Swann v. Charlotte-Mecklenburg Board of Education*, 402 U.S. 1 [1971]). Following this ruling, CM adopted a race-based student busing policy that resulted in many students attending schools that were not located in their own residential neighborhoods. The plan stated that no school was to be more than 50% black and "the burdens of busing" were to be shared equally. To achieve this goal, the plan used noncontiguous satellite zones and the pairing of inner-city black schools with outlying white schools.[7] Since faculties were also segregated by race, teachers were reassigned to schools in 1972 on the basis of their race. After the initial period of reassignment, teacher race was no longer used in the placement or reassignment of teachers to schools.[8] Teachers who were dissatisfied with their schooling assignment in 1972 would have almost three decades to undo any undesirable forcible relocation before the policy change in 2002. As such, any increased reshuffling observed in 2002 can reasonably be attributed to changes in student characteristics.

During the period of student reshuffling between 2002 and 2003, teacher assignment policies remained unchanged. As far back as 1990, the teacher allocation system has operated as follows: teachers in CM can either apply to the school district or apply for an advertised position at a particular school. Advertised positions are those that cannot be easily filled by applicants in the general pool.[9] For advertised positions, applications may be sent to several schools, and the applicant is assigned to the first school that accepts her application. For other openings, principals are provided with a list of eligible applicants, who were selected from the pool of

[7] The plan was subsequently tweaked to accommodate the growth of the black student population and the emergence of magnet schools, but it remained largely the same (legal briefs from *Capacchione v. Charlotte-Mecklenburg Board of Education*, http://www.usdoj.gov/crt/briefs/belk.pdf).

[8] This statement has been verified by the following members of the CM Board of Education: the chief communications officer, lawyers at the CM office of general council, and the director of employee relations. The logic of no longer enforcing the teacher desegregation order was that once students were integrated, teachers could not segregate themselves from students of another race.

[9] For example, teaching positions for kindergarten through grade 3 are often not advertised because they are easy to fill from the existing applicant pool. In contrast, middle school and high school math teacher positions and exceptional children teaching positions are often difficult to staff from the applicant pool and are therefore specifically advertised by the human resources division.

available candidates on the basis of their qualifications and the school's proximity to their home. The district is then notified of the principals' selections from the list, and teachers who are not selected within this group are sent back to the applicant pool to be eligible for other positions. After being assigned, teachers are eligible for a voluntary transfer after having spent 2 years in their current position (unless they wish to move to an understaffed or underperforming school). The transfer application and assignment policy is identical to the application procedure for advertised positions in the district.

In 1997, the CM school system was sued by a parent charging that his daughter was twice denied entrance to a magnet school because the non-black slots were filled and she was not black. This suit was the catalyst for a lengthy legal battle that resulted in the implementation of a neighborhood-based school choice plan for the 2002–3 school year. Under the new policy, students were no longer bused into schools across neighborhoods, and parents listed three schools that they would like their child to attend. If the neighborhood school was the parents' first choice, the student was guaranteed admission. If the parents' most-preferred school was not their neighborhood school, their child would have to enter a lottery in which low-income students were given preference. Those students not admitted to one of their three choice schools were sent to their neighborhood school. Under the new plan, the likelihood that a student would attend a school outside of his or her own neighborhood was significantly reduced.

I use school-level aggregate data from the Common Core of Data available from the National Center of Education Statistics for the years 2000–2005 to determine the impact of this policy change on the demographic makeup of students at CM schools. I augment this data set with school-level achievement and teacher data from the North Carolina Education Research Data Center (NCERDC) and neighborhood (block group)[10] demographic data from the 2000 decennial Census. Since CM is the largest and most urban school district in North Carolina, it is most appropriate to use other large urban school districts as comparison districts. Panel A of table 1 summarizes the school-level student demographic, achievement, and census data for the busing and post-busing years for schools in the CM district, the three next-largest school districts (i.e., "comparison districts": Wake, Guilford, and Cumberland), and all other schools in North Carolina. Of the 152 CM schools in the sample between 2000 and 2005, 137 of them were in operation in 2002. Of these, 86 were primary schools, 29 were middle schools, 15 were high schools, and seven did not fall into any of these categories.

[10] Zip code data are used where block group data are not available.

It is clear that CM was not representative of North Carolina and that CM schools were much more similar to those in the three next-largest school districts. The CM schools were very similar in enrollment to the comparison schools but much larger than other North Carolina schools. CM was the most urbanized district (about 81% of the schools were in a large or midsized city), with the highest share of black students (about 49%) and the lowest share of white residents (about 59%). The comparison schools were somewhat less urbanized (almost 70% of the schools were in a large or midsized city), had lower black enrollment shares (about 41%), and had a higher share of white residents (about 66%). In table 1, one can see that only 27% of schools in the rest of the state were located in a large or midsized city, the average black enrollment share was just over 30%, and whites made up 72% of the residents. The CM schools and those in the comparison districts were located in neighborhoods with median census household incomes of between $46,000 and $49,000 a year, compared to only about $36,000 for schools in the rest of the state. While all schools in the state became increasingly Hispanic during the sample period, there was a somewhat larger increase in CM schools. Across the two time periods, the percentage of students in free-lunch programs increased about 7 points in CM, compared to 4 points in comparison schools and less than 1 point in other schools.

To illustrate the effect of the policy change on the percentage of black students in CM, figure 1 shows kernel density plots of the distribution of the percentage of black students in CM schools and in comparison schools in the years before and after the policy change. Figure 1 illustrates that before the policy change (2000–2002), the distribution of percent black at the schools was relatively similar between CM and the comparison districts. The figure also shows that the distribution became much more dispersed after the policy change (2003–5) in CM, whereas there was almost no change for the comparison districts.

The black differential (BD) variable in table 1 is the percentage of black students at the school in the year 2000 minus the percentage of black residents in the local neighborhood's block group in the year 2000. This variable does not change for a school over time because it is based on data from the year 2000. Schools in both CM and comparison districts were located in areas with about 13 percentage points more black students than the percentage of black residents in the surrounding neighborhoods, compared with 9 percentage points for other schools. This difference may have been due to black families in North Carolina being more likely to have school-age children than white families, or it may reflect the fact that white households were more likely to send their children to private schools. The difference in the gap across school districts could also reflect

Table 1
Summary Statistics for CM, Comparison Districts, and the Rest of North Carolina, by Pre- and Post-policy Change

	Charlotte-Mecklenburg		Comparison Districts		Rest of North Carolina	
	2000–2002	2003–5	2000–2002	2003–5	2000–2002	2003–5
A. School-level variables:						
Black differential (2000)[a]	13.96	13.96	12.55	12.55	8.81	8.81
	(19.96)	(19.96)	(20.11)	(20.11)	(15.98)	(15.98)
School enrollment	762.16	837.66	724.79	727.58	556.57	561.57
	(497.13)	(524.19)	(427.2)	(447.66)	(318.16)	(332.08)
Black students (%)	48.11	49.43	40.78	42.12	30.99	30.58
	(18.93)	(24.49)	(21.61)	(22.32)	(26.5)	(25.97)
White students (%)	40.89	35.61	50.57	46.95	61.79	59.95
	(21.59)	(26.95)	(22.33)	(22.62)	(27.88)	(27.89)
Hispanic students (%)	6.29	10.21	4.64	6.66	4.25	6.40
	(6.61)	(9.59)	(3.99)	(4.83)	(5.55)	(7.73)
Asian students (%)	4.11	4.07	3.18	3.45	1.14	1.23
	(2.41)	(2.45)	(3.21)	(3.55)	(2.35)	(2.38)
Free-lunch-eligible students (%)	38.01	44.90	32.05	36.02	37.29	37.63
	(21.35)	(27.4)	(20.43)	(21.67)	(20.56)	(23.49)
Median household income (2000 Census)	48,366	48,272	47,225	46,868	35,993	36,031
	(15,612)	(15,774)	(15,805)	(15,598)	(8,068)	(8,154)
Black residents (2000 Census) (%)	35.30	35.30	28.03	28.03	22.69	22.69
	(23.69)	(23.69)	(20.95)	(20.95)	(18.71)	(18.71)
White residents (2000 Census) (%)	59.01	58.44	66.13	65.79	72.41	72.41
	(23.15)	(23.81)	(20.99)	(20.98)	(20.2)	(20.26)
City	.80	.82	.66	.72	.26	.27
	(.4)	(.39)	(.47)	(.45)	(.44)	(.44)
At or above grade level:						
Math (%)	78.01	85.30	83.05	87.48	80.56	86.32
	(13.42)	(12.31)	(12.53)	(10.74)	(14.39)	(11.67)
Reading (%)	72.80	79.39	79.26	83.43	76.08	81.84
	(14.53)	(12.95)	(13.2)	(11.36)	(14.14)	(11.11)

B. Teacher variables:						
0–3 years' experience (%)	32.06	30.99	25.42	25.72	22.44	21.45
	(12.1)	(11.84)	(11.88)	(10.7)	(11.)	(10.39)
4–10 years' experience (%)	27.33	30.36	26.49	27.60	24.96	26.22
	(7.68)	(8.17)	(9.88)	(8.78)	(9.01)	(8.77)
11+ years' experience (%)	40.61	38.65	48.09	46.69	52.60	52.27
	(11.42)	(12.37)	(13.41)	(12.33)	(12.9)	(12.57)
1-year teacher turnover rate[b]	27.65	25.23	24.94	22.85	21.39	18.98
	(13.26)	(13.06)	(10.94)	(10.41)	(11.11)	(10.21)
Black (%)	23.78	24.57	20.91	23.44	13.41	13.45
	(15.47)	(17.56)	(17.21)	(18.14)	(17.22)	(18.25)
White (%)	74.40	72.40	77.12	73.49	84.66	84.39
	(15.81)	(18.22)	(17.64)	(19.04)	(18.61)	(19.66)
Advanced degree (%)	19.59	21.66	17.37	18.10	11.70	11.94
	(14.43)	(15.09)	(12.02)	(12.81)	(8.36)	(8.7)
Score in top 25% (%)	47.12	47.86	46.59	48.55	42.56	45.24
	(10.12)	(11.18)	(12.67)	(13.43)	(13.91)	(14.22)
Score in top 50% (%)	73.28	75.55	71.92	74.39	69.76	71.54
	(9.97)	(9.67)	(12.44)	(12.39)	(13.83)	(13.66)
Top-100 college (%)	9.06	12.80	12.87	15.46	7.94	10.00
	(5.4)	(6.33)	(10.58)	(11.31)	(7.45)	(8.76)
Number of schools	152		358		2,220	

NOTE.—Standard deviations are in parentheses. The unit of observation is a school year. Each school has one observation in each year in the sample. Since the panel is not balanced because of new schools or school closings, variables that do not vary over time may change on average across time because of composition effects. The comparison districts are Wake, Guilford, and Cumberland.

[a] Black differential is defined as the percentage of black students at the school in the year 2000 minus the percentage of black residents in the census block group (or zip code if black group data are not available) of the school in the 2000 Census. In Charlotte-Mecklenburg this variable ranges from −31.34 to +57.06.

[b] The teacher turnover rate is computed in the sample so that errors in data classification or missing data would inflate teacher turnover. This should not affect regression results that are based on changes in this variable.

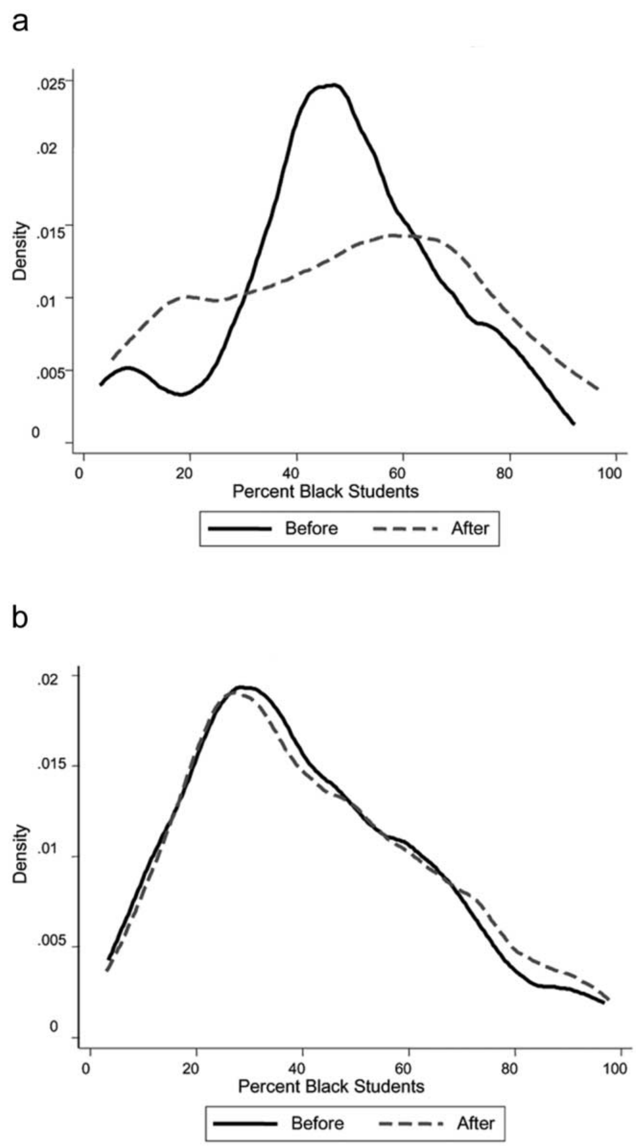

Fig. 1.—Change in the distribution of percent black students in (*a*) CM and (*b*) comparison schools before and after the 2002 policy change. The comparison school districts are Wake, Guilford, and Cumberland.

a greater number of white students in urban environments going to private schools.[11]

Since the busing policy that ended in 2002 maintained school integration despite much residential segregation, the schools that would be expected to have experienced the greatest change in student demographics are those that had proportionately more blacks or whites in the school than in the surrounding area.[12] A school with 10% black students located in an area with 50% black residents (a BD of −40) would have a larger inflow of black students at the end of busing than a school with 90% black students in a neighborhood in which 100% of the residents were black (a BD of −10). The BD predicts the outflow of black students that would have occurred if all schools had had student populations that were exactly representative of the surrounding neighborhoods. A variable denoting post-busing, equal to one after 2002 and zero otherwise, would identify the year in which schools were most likely to have student populations that mirrored the attributes of the surrounding neighborhoods. By interacting the BD variable with a "post" variable, one can predict the exogenous change in the share of black students that was due to the policy change. To illustrate this point, figure 2 shows the relationship between the BD of a school in 2000 and the change in the percentage of black students between 2001 and 2002 (the year before the policy change) and between 2002 and 2003 (the year of the policy change).

The two left scatter plots show the sizable difference in the relationship between BD and changes in the percentage of black students before and after the 2002 policy change in Charlotte-Mecklenburg. As one would expect, the two right scatter plots show very little difference over time for the comparison districts. The BD predicts small changes in the percentage of black students in CM schools in the pre-policy year and in the comparison districts for all years, such that schools with negative BDs (fewer blacks than predicted by the makeup of the neighborhood) ex-

[11] Even though the comparison districts did not have student busing during the sample period, they all did in the past so that old district lines still crossed neighborhoods, where possible, to maintain diversity within schools. Wake County, the second-largest county in North Carolina, moved from a race-based to an income-based busing system in 2000, so there still were forces keeping BD high in Wake. While Cumberland and Guilford counties did not have student busing policies, they both explicitly aimed to maintain racial diversity across school districts when drawing and redrawing school enrollment areas.

[12] In regressions that predict the change in the percentage of black students in schools, the difference between the percentage of black students in the school and the percentage of black residents in the neighborhood has a much larger *F*-statistic than simply using the percentage of black residents in the neighborhood.

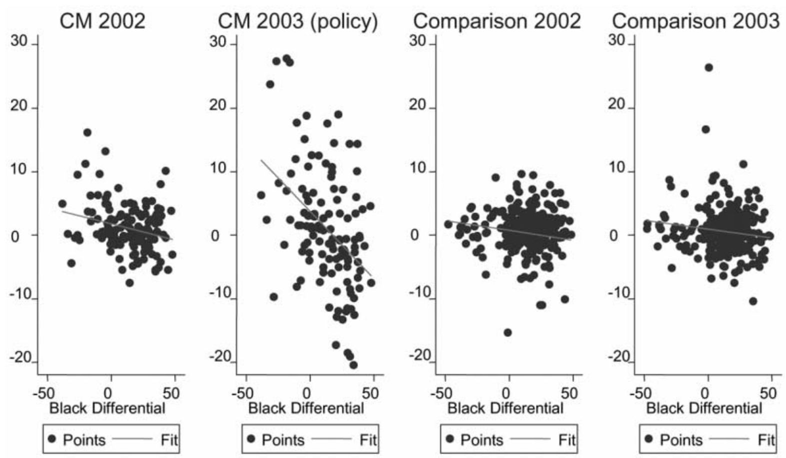

FIG. 2.—Relationship between black differential and changes in percent black students, by district and year. The comparison school districts are Wake, Guilford, and Cumberland. The *y*-axis shows the 1-year change in the percentage of black students at the school.

perienced small increases in the share of black students.[13] In contrast, BD predicts large changes in the percentage of black students in CM during the policy change year (2002–3). Also illustrated in figure 2 are the mechanics of the instrument that uses the difference in the change in the relationship between BD and the percentage of black students at the school before and after the policy change between CM and the comparison schools. Most schools that experienced large inflows of black or white students between 2002 and 2003 were located in predominantly white or black neighborhoods, respectively. Therefore, the instrument predicts the local average treatment effect—the effect of an inflow or outflow of black students on schools in largely black or white neighborhoods, respectively.

To demonstrate further that the BD variable predicts a sudden inflow or outflow of black students between 2002 and 2003 above and beyond that in other years, I estimate the within-school change from 1998 levels in the proportion of black students for the CM schools with BDs above the 75th percentile and for those with BDs below the 25th percentile. Figure 3 shows that schools with BDs above the 75th percentile (predictive of an outflow of blacks) experienced a slight decrease in the share of black students between 2002 and 2003, whereas schools with BDs below the

[13] Note that one cannot reject the hypothesis that the relationship between BD and within-school changes in the percentage of black students between 2001 and 2002 is the same in CM as in the comparison districts (at traditional levels). This indicates that the comparison districts may provide credible counterfactual changes.

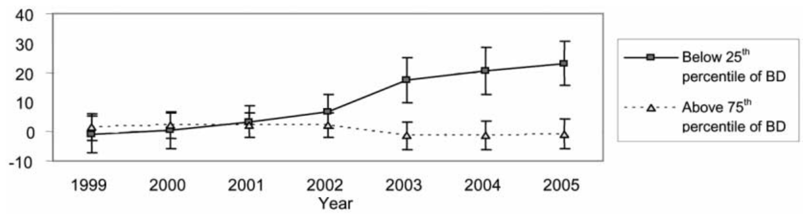

Fig. 3.—Change in the percentage of black students by black differential percentile group (relative to 1998 levels). The figure plots the estimated year fixed effects from a linear regression that includes individual school indicator variables on the two subsamples of schools (those schools below the 25th percentile and those above the 75th percentile of the black differential distribution).

25th percentile (predictive of an inflow of blacks) experienced an increase in the share of black students during the same time period. The figure suggests that BD predicts relatively sudden *differential* changes in the share of black students during the year of the policy change.[14]

The effect of the policy change on student characteristics.—To describe the change in student characteristics to which teachers were exposed (i.e., the treatment), I ran regressions to determine the effect of the policy change on various student characteristics. While the final analysis uses the percentage of black students as the treatment, teachers are exposed to all student characteristics that are associated with black students. Therefore, it is instructive to look at other student characteristics. It is useful to consider the student demographics first-stage regressions in which the coefficients on the instruments predict the treatment that schools and teachers are exposed to.

Since the policy change had a differential effect on high-BD versus low-BD schools, one could, in principle, identify the effect of the policy change using a difference in difference (DID) estimator, comparing the difference in the change in outcomes between 2002 and 2003 across high-BD and low-BD schools in CM. This DID strategy would be valid if high-BD schools and low-BD schools would have experienced the same change in outcomes in the absence of the policy change. Since high-BD and low-BD schools were located in different neighborhoods and served different populations, the assumption that they had the same underlying dynamics is implausible. In addition, statewide policies aimed at particular types of schools (e.g., low-income, low-performing) may have had differential effects across school types and would invalidate the exclusion restriction

[14] In addition to the movement of students across public schools, the change in the share of black students could also reflect the movement of white students from private schools back into public schools that had lost a large fraction of black students in 2003.

in a standard DID approach. For example, the North Carolina Bonus Program, which paid teachers for locating in low-performing schools, was implemented in 2001 and differentially affected teacher turnover across high- and low-income schools in 2002.

To address this concern, I use schools from the three next-largest school districts (Guilford, Wake, and Cumberland) as comparison schools, allowing me to introduce another round of differencing and to implement a difference in difference in differences (DIDID) estimator.[15] Identification in this triple-differenced model compares the difference in the change in outcomes between high-BD and low-BD schools within CM (which had the policy change) to that of other school districts (which did not have the policy change). The identifying assumption is that the difference in the change in outcomes between high-BD and low-BD schools in the comparison districts is the difference in the change in outcomes that would have occurred in CM between high-BD and low-BD schools had there been no policy change. Figures 1 and 2 suggest that this assumption is plausible. This assumption is more compelling than that for the standard DID approach since the DIDID approach will "net out" any statewide policies or differential migration that could have had a different time effect across different types of schools. Since the predictor for an inflow of black students (BD) is computed on the basis of data in 2000, the estimation sample does not include data before 2000 to avoid any mechanical endogeneity between the instrument and the variables of interest.[16] The first set of basic DIDID estimates are implemented by estimating the following equation by ordinary least squares (OLS) on the schools in the four largest school districts in the state (all subsequent analyses are based on this sample of schools):

$$\begin{aligned} Y_{it} = {} & \delta \cdot \mathrm{POST}_t \times \mathrm{CM}_i \times \mathrm{BD}_i + \omega_1 \mathrm{POST}_t + \omega_2 \mathrm{POST}_t \times \mathrm{BD}_i \\ & + \omega_3 \mathrm{POST}_t \times \mathrm{CM}_i + \theta_i + \varepsilon_{it}. \end{aligned} \tag{1}$$

In (1) Y_{it} is the outcome for school i at time t, POST_t is an indicator variable equal to one in the year 2003 onward and zero otherwise,

[15] There is an efficiency/consistency trade-off in increasing the sample to all schools in North Carolina. Since CM is the largest and most urbanized school district in the state, restricting the comparison sample to other large, urban school districts is desirable. I chose the four largest school districts because the size and urbanicity of school districts change rather suddenly as one goes beyond the first few largest districts. For example, student enrollments for the year 2000 for the four largest districts were 103,000, 99,000, 63,000, and 51,000. For the next three largest districts the enrollments were 44,000, 30,000, and 30,000. Restricting the analysis to the top three districts results in less power but does not change the results in any meaningful way.

[16] Excluding data for the year 2000 is unnecessary since all regression specifications are differenced.

CM_i is an indicator variable equal to one if school i is in Charlotte-Mecklenburg and equal to zero otherwise, BD_i is the black differential for school i, θ_i is a school-specific intercept, and ε_{it} is the idiosyncratic error term. The school dummies θ_i subsume the necessary one-way and two-way effects between CM and BD. The parameter of interest is δ, the coefficient on the three-way interaction $\mathrm{POST}_t \times \mathrm{CM}_i \times \mathrm{BD}_i$ that predicts an outflow of black students.

The regression results in table 2 show that the policy had a strong effect on the racial composition of students at the affected schools and that CM schools with more black students than the neighborhood demographics would predict experienced an outflow of black students and an inflow of white students between 2002 and 2003 after busing ended. Specifically, the -0.253 coefficient for the variable $\mathrm{POST}_t \times \mathrm{CM}_i \times \mathrm{BD}_i$ in column 1 indicates that a school in CM would have had a $0.253 \times 20 = 5.06$-point greater increase in the percentage of black students between 2002 and 2003 than a school in CM over the same time period with a black differential 20 points higher (a one standard deviation difference in BD). The t-statistic on the coefficient is 4.77, indicating a strong first stage. The odd-numbered columns show that relative to a school in CM with a BD of zero, a school in CM with a BD of 20 would have had a 5.06-point decrease in the percentage of black students, a 3.6-point increase in the percentage of white students, and a 6.16-point decrease in the percentage of students who were in a free-lunch program. CM schools also experienced changes in student achievement. Note that changes in achievement could also reflect the effect of teacher mobility, peer quality, or other unmeasured inputs that may be endogenous to student race rather than simply changes in the ability of students. A school in CM with a black differential of 20 would have experienced a 2.08- and 3.34-point increase in the percentage of third through eighth grade students at or above grade level in math and reading, respectively, relative to a CM school with a BD of zero over the same time period.

While the DIDID specification is instructive, I augment model (1) to control for neighborhood characteristics and to allow for a more flexible specification. Specifically, I include year effects instead of a simple before/after dummy, use district fixed effects instead of a simple CM dummy, and include neighborhood characteristics interacted with year fixed effects. More formally, I estimate equation (2) below by OLS:

$$\begin{aligned} Y_{it} = {} & \delta \cdot \mathrm{POST}_t \times \mathrm{CM}_i \times \mathrm{BD}_i + \omega_2 \mathrm{POST}_t \times \mathrm{BD}_i \\ & + \omega_{3,r} \sum_r^6 I_{\mathrm{year}=r} \times \mathrm{LOC}_i + \omega_{4,r} \sum_r^6 I_{\mathrm{year}=r} \times \mathrm{DEC}_i \\ & + \omega_{5,r} \sum_r^6 I_{\mathrm{year}=r} \times \mathrm{DISTRICT}_i + \theta_i + \varepsilon_{it}. \end{aligned} \quad (2)$$

Table 2
The Effect of the Policy Change on School Attributes: OLS Estimation

	Black Students (%)		White Students (%)		Student Free-Lunch Eligible (%)		At Grade Level (Math) (%)		At Grade Level (Reading) (%)	
	(1)[a]	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
POST × CM × BD	−.253	−.2431	.18	.1815	−.308	−.2156	.104	.154	.167	.1965
	[.053]***	[.0480]***	[.054]***	[.0488]***	[.079]***	[.0815]***	[.050]***	[.0504]***	[.056]***	[.0583]***
POST × CM	5.249	. . .	−5.896	. . .	9.103	. . .	−.206	. . .	−1.911	. . .
	[1.349]***	. . .	[1.397]***	. . .	[2.064]***	. . .	[1.294]	. . .	[1.482]	. . .
POST	2.485	. . .	−4.942	. . .	4.689	. . .	4.259	. . .	4.114	. . .
	[.391]***	. . .	[.585]***	. . .	[.599]***	. . .	[.616]***	. . .	[.563]***	. . .
POST × BD	−.061	. . .	.059	. . .	.026	. . .	.031	. . .	.024	. . .
	[.026]**	. . .	[.025]**	. . .	[.031]	. . .	[.018]*	. . .	[.015]	. . .
School dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Locale × year dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Census % black residents decile × year dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
District × year dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,542	2,503	2,503	2,503	2,514	2,475	1,801	1,777	1,801	1,777
Number of schools	431	419	419	419	431	419	370	358	370	358
R^2	.22	.31	.35	.45	.2	.32	.3	.42	.3	.42

Note.—The dependent variable is above each column. Robust standard errors are in brackets. Robust standard errors are clustered at the zip code level. The sample is CM, Wake, Guilford, and Cumberland districts. BD is the percentage of black students at the school (in 2000) minus the percentage of black residents in the block group or zip code (in 2000) in which the school is located. CM stands for Charlotte-Mecklenburg and POST denotes after the policy change (2003 onward). The POST × CM × BD variable is a difference-in-difference-in-difference estimate.

[a] First stage.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

All common variables are defined as in (1), and $I_{\text{year}=r}$ is an indicator variable equal to one if the observation year is year r and zero otherwise. To control for underlying dynamics that may have had a differential effect across school districts, urban environments, and neighborhoods with different shares of black residents, I include interactions of $I_{\text{year}=r}$ dummies with indicators for each school district, DISTRICT_i; indicators for the urbanity of the surrounding area, LOC_i; and dummies denoting the 10 deciles of the distribution of the percentage of black residents in the surrounding area, DEC_i. The results of this more flexible model are presented in the even-numbered columns of table 2. The flexible specification yields results similar to those of equation (1) and is used for all subsequent analysis.

In sum, table 2 shows that the student body changed in a variety of ways associated with student race, such as income levels and achievement levels. For the remainder of the paper, I use the change in the percentage of black students to categorize the change in student demographics. As such, the results on teacher characteristics must not be interpreted as being the result of teachers having preferences for student race per se, but the result of teachers having preferences for student or school characteristics that are endogenous to student race—such as student achievement, school culture, student behaviors, and parental characteristics.

IV. The Effect of the Policy Change on Teacher Characteristics

In this section I analyze aggregate teacher data to determine the effect of the policy change on teacher attributes. The teacher data were created by computing school-level aggregate statistics from individual teacher data from the NCERDC. The rankings of the colleges or universities teachers attended were obtained by linking *U.S. News and World Report* rankings from 2005 to the undergraduate institution data from the teacher education files. Teacher license score data were created by comparing each teacher's score on the exam to that of all other teachers in the state in that year. Variables were created denoting whether the teacher scored above the 75th percentile or the median on that exam in that year. Since teachers may have taken more than one exam, I code a teacher as having scored above the 75th percentile or the median if she has at least one score above the 75th percentile or the median on any one exam. Therefore, more than half of the teachers would be expected to score above the median. Teacher value added was computed by linking the student end-of-year test files with individual teacher data. Since teacher effectiveness could have been affected by changes that took place as a result of students' demographics changing or teacher demographics changing between 2002 and 2003, teacher value added is estimated "out of sample" for the years 1995–2000.

Although there are several specifications used in the literature to estimate teacher value added, the estimated teacher fixed effects across studies are surprisingly robust to the chosen specification.[17] To identify effective teachers, I estimate teacher fixed effects in a test score growth model of the form (3) using data on students in grades 3–5 from 1995 through 2000:[18]

$$A_{ijgt} - A_{ijg-1t-1} = \psi_1 A_{ijg-1t-1} + \psi 2\bar{A}_{i'jg-1t-1}\psi_3 X_i + \psi_4 Z_{st} + \psi_5 W_{jt} + \tau_j + \tau_t + \tau_g + v_{jt} + \varepsilon_{ijgt}. \quad (3)$$

In (3) A_{ijgt} is the achievement score of student i with teacher j in grade g in year t; $\bar{A}_{i'jg-1t-1}$ are the average incoming test scores of a student's classmates; and X_i is a vector of student characteristics such as ethnicity, gender, and parental education level. The term W_{jt} is a vector including teacher experience, class size, and variables denoting the gender and ethnic match between the student and the teacher;[19] Z_{st} is a vector of school attributes including the percent black, percent white, percent Hispanic, percent free-lunch-eligible students, and urbanicity of the school (whether the school is in a large city, medium-sized city, urban fringe, suburban area, or rural area); τ_t is a year fixed effect; τ_g is a grade fixed effect; τ_j is a teacher effect; v_{jt} is a classroom-level error term; and ε_{ijgt} is the idiosyncratic student-level error term. Since I need estimates of teacher value added that are comparable across schools, grades, and classes, I do not include school or student fixed effects but rather include a set of de-

[17] For a detailed discussion of the theoretical and econometric assumptions underlying value-added specifications, see Todd and Wolpin (2003).

[18] Researchers have pointed out that measurement error in the lagged test score could bias estimates of the coefficient of lagged test scores on test score growth. The common fix for this problem is to assume that there is no serial correlation in the error terms over time and (where there are enough data) to instrument for lagged test scores with the second lag of test scores. The main results do not use this approach since it results in a small estimation sample that makes identification of teacher fixed effects difficult. (I lose one additional year of data to include the second lag, resulting in an estimation sample of 3 years.) I do, however, present results in app. table B2 showing that making this correction yields results similar to those of the chosen specification despite producing noisier estimates.

[19] The value-added results are robust to omitting the gender and ethnic match variables.

mographic controls for the students and schools.[20] Readers may be concerned that the included covariates do not adequately capture measures of school quality, in which case the teacher effects would capture school, principal, and other unobserved effects.[21] Although this is possible, these estimates are used in a within-school model on an out-of-sample period so that changes in the distribution of these estimates within schools over time will not be confounded with those unobservable school inputs. The estimates of regression equation (3) are in appendix table B1. The teacher value-added estimates τ_j are standardized, normalized, and linked to teachers in the 2000–2005 data, and school-level aggregates are computed. I also compute shrinkage estimates, or empirical Bayes (EB) estimates, that shrink noisy teacher value-added estimates toward zero for greater statistical precision. Details of how the EB estimates are constructed are in appendix A. Results using normalized teacher estimates directly from the regression are similar to those using the normalized EB estimates.

It should be noted that not all teachers have estimated teacher effects since not all teachers teach basic English and math, and teachers who were not in the sample in 2000 would not have estimated teacher value added. As such, changes in the distribution of estimated teacher value added within schools over time reflect changes in the distribution of those teachers who were in the sample in the year 2000, but not necessarily in the distribution of new teachers or teachers who had experience but came from outside of North Carolina. Also, since I estimate teacher value added for teachers in primary school between 1995 and 2000, the school-level aggregates are defined for schools that employed primary school teachers after 2000. Note that (1) primary schools make up about two-thirds of all schools in CM, and (2) the findings are robust to looking at changes in teacher value added for primary schools only.

Panel B of table 1 summarizes the teacher variables used. Teacher turn-

[20] Specifications that include student or school fixed effects identify teacher value added on the basis of within-school or within-student variation. If teachers are very different across schools, then much of the variation in teacher quality (i.e., the cross-school variation) will be absorbed by the school fixed effect, making estimated effects across schools impossible to compare. Including student fixed effects further exacerbates this problem by allowing comparisons only of teachers who teach the same groups of students. If those teachers who teach the gifted and talented students are of different average quality than those who teach the regular students, the estimated teacher value added can be used to compare only teachers who share the same students, so comparing teachers who teach different students (even within the same school) may be misguided.

[21] Note that using within-school or within-student variation to identify teacher value added loads any common effectiveness at a school on the school even if benefits are due to the teachers. Such models also lead to attenuated teacher effects if there are spillovers across teachers. However, results using student fixed effects are qualitatively similar to those presented here.

over was somewhat higher in the large school districts than in the rest of the state (about 26% for CM and 23% for the three comparison districts, compared to about 20% for the rest of the state). Consistent with this, CM and the comparison districts had larger shares of rookie teachers and lower shares of experienced teachers. These districts also had a greater share of black teachers (about 24% for CM and 22% for the three comparison districts, compared to about 13.5% for the rest of the state), a greater share of teachers with advanced degrees, and a greater share of teachers who attended a top-100 college than other schools in the state did.

To determine whether the change in student demographics affected schools' overall teacher makeup, I run regressions of teacher characteristics on the percentage of black students. To use only variation in black enrollment shares that are attributable to the policy change, I instrument for the percentage of black students with the triple-differenced $\text{POST}_t \times \text{CM}_i \times \text{BD}_i$ variable from equation (2). Specifically, I estimate the following system of equations by two-stage least squares (2SLS):

$$\begin{aligned}\%\text{black} = {} & \pi_1 \cdot \text{POST}_t \times \text{CM}_i \times \text{BD}_i + \pi_2 \text{POST}_t \times \text{BD}_i \\ & + \pi_{3,r} \sum_r^6 I_{\text{year}=r} \times \text{LOC}_i + \pi_{4,r} \sum_r^6 I_{\text{year}=r} \times \text{DEC}_i \\ & + \pi_{5,r} \sum_r^6 I_{\text{year}=r} \times \text{DISTRICT}_i + \theta_{1i} + \varepsilon_{1it}, \end{aligned} \tag{4}$$

$$\begin{aligned} Y_{it} = {} & \delta_2 \cdot (\%\text{black}_{it}) + \phi_2 \text{POST}_t \times \text{BD}_i + \phi_{3,r} \sum_r^6 I_{\text{year}=r} \times \text{LOC}_i \\ & + \phi_{4,r} \sum_r^6 I_{\text{year}=r} \times \text{DEC}_i + \phi_{5,r} \sum_r^6 I_{\text{year}=r} \times \text{DISTRICT}_i \\ & + \theta_{2i} + \varepsilon_{2it}. \end{aligned} \tag{5}$$

All variables are defined as in (2), and equation (4) is equation (2) with percent black as the dependent variable shown in column 2 of table 2. In the second-stage regression, the fitted values from (4) are used in place of percent black_{it} in (5).[22] The excluded instrument in (5) is the three-way interaction $\text{POST}_t \times \text{CM}_i \times \text{BD}_i$, and Y_{it} is the teacher outcome for school i at time t. Since the model includes year effects by district, locale, and decile of the percentage of black residents in the neighborhood, the pa-

[22] I estimate eqq. (4) and (5) by 2SLS using the xtivreg2 command in STATA. This command automatically adjusts the standard errors in the second stage for estimation error and uses the appropriate degrees of freedom adjustment.

rameter δ_2 identifies the effect of an inflow of black students that is arguably uncorrelated with those changes that may have naturally occurred across different neighborhoods over time.

To highlight the differences between the cross-sectional relationships and the relationships one observes based on the policy change, I also estimate a simple model of the outcome of interest on percent black and a constant (OLS regression). Table 3 documents the cross-sectional relationship between the percentage of black students at a school and various teacher characteristics in column 1. Column 2 presents results of an intermediate DIDID specification, and the instrumental variables DIDID (IV-DIDID) regression results are reported in column 3. Table 3 shows the coefficient on percent black for each outcome and each model.

The standard deviation of the change in percent black in CM between 2002 and 2003 is just over 10. This is also approximately the amount of variation associated with a two standard deviation difference in BD. Column 1 shows that in the cross section, a school with 10 percentage points more black students would have 1.53 percentage points more teachers with 0–3 years of experience, a teacher turnover rate 1.86 points higher, 5.26 percentage points more black teachers, 5.28 percentage points fewer white teachers, 0.73 percentage points fewer teachers with an advanced degree, 0.86 percentage points fewer teachers who attended a college ranked in the top 100, approximately 2 percentage points fewer teachers who scored above the 75th percentile and the median on their certification exams, and about 0.04 and 0.02 standard deviations lower mean teacher value added in math and reading, respectively. In sum, schools with large black enrollment shares had teachers with weaker observable characteristics on average.

Column 2 shows an intermediate specification documenting the relationship between changes in student demographics within schools over time and changes in teacher characteristics so that the reader may see the marginal effect of going to the IV model. As one can see, although the estimated coefficients in column 2 are smaller than those in column 1, the results are qualitatively similar. Column 3 documents the relationship between student demographics and teacher characteristics using the variation that is due to the policy change. The IV-DIDID estimates show that a 10-point increase in the percentage of black students due to the policy change is associated with a decrease of 0.8 years in the average experience of teachers at the school. This is much larger than the OLS and DIDID estimates of only 0.27 and 0.32 years, respectively. Rows 2–5 indicate that this change is due to an increase in the share of teachers with fewer than 10 years of experience and a decrease in the share of teachers with 10 or more years of experience.

Row 6 shows the surprising result that schools that had an inflow of black students did not experience a greater increase in turnover than

Table 3
The Effect of the Percentage of Black Students on Teacher Characteristics

Dependent Variable	OLS (1)		DIDID (2)		DIDID-IV[a] (3)		DIDID-IV[b] (4)	
1. Teacher experience (mean)	−.028	[.006]***	−.0324	[.0096]***	−.087	[.038]**	−.088	[.032]***
2. Teachers: 1–3 years (%)	.153	[.020]***	.1036	[.0460]***	.163	[.185]	.256	[.1390]*
3. Teachers: 4–9 years (%)	−.006	[.018]	.0395	[.0426]	.184	[.099]*	.0982	[.0833]
4. Teachers: 10–20 years (%)	−.096	[.013]***	−.0654	[.0361]*	−.156	[.125]	−.1526	[.1108]
5. Teachers: 21 years (%)	−.051	[.018]***	−.0777	[.0318]**	−.191	[.083]**	−.2016	[.0718]***
6. Teachers: leave current school (%)	.186	[.023]***	.1051	[.0639]	−.123	[.174]	−.0907	[.1588]
7. Lag teachers leave current school (%)	.167	[.020]***	.0561	[.0680]	−.342	[.230]	−.1394	[.1514]
8. Teachers: black (%)	.526	[.039]***	.2165	[.0510]***	.373	[.159]**	.3566	[.1618]**
9. Teachers: white (%)	−.528	[.038]***	−.1864	[.0503]***	−.299	[.178]*	−.2963	[.1801]*
10. Teachers: higher degree (%)	−.073	[.019]***	−.0681	[.0294]**	.077	[.145]	−.0722	[.1242]
11. Teachers: top-50 college (%)	−.05	[.008]***	−.0172	[.0174]	.006	[.047]	−.0065	[.0542]
12. Teachers: top-100 college (%)	−.086	[.013]***	−.0328	[.0240]	−.011	[.071]	−.0327	[.0741]
13. Teachers: top-10 score (%)	−.133	[.018]***	−.0873	[.0414]*	−.13	[.163]	−.2081	[.1215]*
14. Teachers: top-25 score (%)	−.205	[.019]***	−.1538	[.0607]**	−.177	[.228]	−.2556	[.1947]
15. Teachers: top-50 score (%)	−.211	[.022]***	−.082	[.0502]	−.099	[.159]	−.1465	[.1325]
16. Teacher value added math (mean)	−.002	[.001]*	−.0003	[.0025]	−.015	[.005]***	−.0134	[.0045]***

17. Teacher value added math (EB) (mean)	−.004 [.001]***	−.0034 [.0022]	−.0206 [.005]**	−.0197 [.0049]**
18. Teacher value added reading (mean)	.0002 [.001]	−.0029 [.0017]*	−.013 [.007]*	−.013 [.0063]*
19. Teacher value added reading (EB) (mean)	−.0021 [.001]**	−.0053 [.0021]**	−.0224 [.006]**	−.0224 [.0057]***
School effects	No	Yes	Yes	Yes
Year-by-district effects	No	Yes	Yes	Yes
Percent black residents decile-by-years effects	No	Yes	Yes	Yes
Locale-by-year effects	No	Yes	Yes	Yes
Excluded instrument	. . .	. . .	POST × CM × BD	(POST × CM × BD) × Q_i

Note.—The coefficient on the percent black is reported. Robust standard errors are in brackets. Standard errors are clustered at the zip code level. The coefficient on the percent black students at the school is reported. Percent black ranges from zero to 100. Each column-row combination represents a different regression. BD is the percentage of black students at the school (in 2000) minus the percentage of black residents in the block group or zip code (in 2000) in which the school is located. CM stands for Charlotte-Mecklenburg and POST denotes after the policy change (2003 onward). The POST × CM × BD variable is a difference-in-difference-in-difference estimate. The term Q_i denotes which of the five quintiles of the distribution of the percentage of black residents the school falls into. The sample is CM, Wake, Guilford, and Cumberland districts (2,503 observations and 419 schools).

[a] The excluded instrument in col. 3 is the black differential of the school interacted with a dummy denoting CM district interacted with a dummy variable denoting after 2002.

[b] The excluded instruments in col. 4 are the interactions of CM × BD × POST with indicator variables denoting the five quintiles of the distribution of the percentage of black residents in the surrounding neighborhood.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

schools that had an outflow. Although schools with larger black enrollment shares had higher teacher turnover in the cross section, this relationship does not hold in the IV results (in fact, the point estimate in col. 3 is negative and not statistically significant). Since there was a period after which teachers would have known about the policy change but before students were actually moved, I also include the 1-year lag of turnover as a dependent variable. There was no statistically significant relationship between lagged turnover and an inflow of black students, and the point estimate is negative. The graphical analysis of teacher turnover in Section V puts this surprising result in perspective.

Rows 8 and 9 show that the relationship between teacher race and student race is robust across specifications. However, the IV estimates indicate that a 10-point increase in the black enrollment share is associated with a 3.5-point increase in the black teacher share compared to a 5.3-point increase in the OLS estimates. The IV-DIDID coefficient is about two-thirds as large as the OLS coefficient, suggesting that much of the correlation between teacher race and student race is an artifact of residential segregation. The fact that there is still a strong relationship in the IV-DIDID results strongly suggests that the relationship between teacher race and student race is not simply an artifact of colocation due to residential segregation, but is due to something systematic about how teachers apply to or are placed in schools. Since there was no change in teacher placement policy and race was not explicitly used in the teacher hiring or placement procedure, it is reasonable to interpret this as a teacher labor supply response.

The results in column 3 show no systematic relationship between the percentage of black students and the percentage of teachers with an advanced degree or the percentage of teachers who attended top-50 or top-100 colleges. The point estimates have the opposite sign of the OLS and intermediate DIDID estimates. The point estimates in rows 13–15 suggest that an inflow of black students is associated with teachers with lower scores on their certification exams, but these estimates are not statistically significant at traditional levels. Rows 16–19 document the relationship between estimated teacher value added (based on a presample period) and the percentage of black students at the school. The IV-DIDID results indicate that a 10-point increase in the share of black students is associated with a 0.15 and 0.13 standard deviation decrease in the average teacher value added in math and reading, respectively. With the EB teacher effects (rows 17 and 19), a 10-point increase in the share of black students is associated with a 0.21 and 0.22 standard deviation decrease in the average teacher value added in math and reading, respectively. These effects are much larger than those from the OLS and the intermediate DIDID specifications.

In column 4, I interact the $\mathrm{POST}_t \times \mathrm{CM}_i \times \mathrm{BD}_i$ variable with Q_i (the

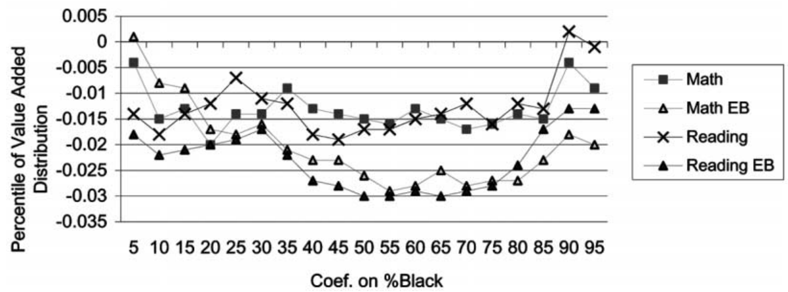

FIG. 4.—Coefficient on percent black students on percentiles of teacher value-added distribution

quintile of the school in the distribution of the percentage of black residents in the neighborhood) to allow the instrument to have a differential effect on schools located in largely black neighborhoods as opposed to largely white neighborhoods.[23] Figure 3 indicates that this is likely to improve the fit of the first stage and reduce noise in the second stage. Making this adjustment to the excluded instrument reduces the standard errors on most estimates. The results are largely the same as those of column 3 of table 3. However, in column 4 of table 3, an increase in the share of black students is associated with a decrease in the share of teachers who scored in the top 10% on the certification exam. This relationship is significant at the 10% level. In column 4, even those outcomes that are not statistically significant have the expected sign and tell the same consistent story: schools that had an exogenous increase in the black enrollment share experienced a decrease in the observable and unobservable quality of teachers on average.

To provide a more nuanced picture of how the distribution of estimated teacher value added changed within schools as a result of the policy change, figure 4 shows the marginal effects of an inflow of black students on different percentiles of the value-added distribution for reading and math. The regression coefficients are reported in appendix table B2. Whether one uses EB estimates or the estimated teacher effects, the results are qualitatively the same: an increase in the share of black students is

[23] There are several reasons why BD may be a stronger predictor of changes in percent black in neighborhoods with more or fewer black residents. For example, if inflows of black students were more likely to induce private school attendance among whites in areas that already had a critical mass of black students, BD would be a stronger predictor of inflows of black students in black neighborhoods than in white neighborhoods.

associated with a statistically significant decrease in the value added of teachers at the school at all points in the value-added distribution.[24]

To put these results into perspective, consider the following "back-of-the-envelope" calculation. Assume that under student busing the average black/white student attended a school that was 60% black/white and after busing attended a school that was 75% black/white. Then they would have been faced with teachers who had, on average, approximately 0.3 standard deviations lower/higher value added in math and reading. This ignores any preexisting differences that might have existed across schools during busing. This would imply an increased teacher quality gap of about 0.6 standard deviations, which would imply an increased performance gap of 7.5% and 3.3% of a standard deviation in math and reading, respectively.[25] This is roughly the magnitude of having a first-year teacher as opposed to a more experienced teacher. The estimated black-white test score gap in CM was about one standard deviation in 2001 in both math and reading. This suggests that the endogenous sorting of teachers with respect to student race could potentially explain between 3.3% and 7.5% of the black-white test score gap in CM.

V. A Graphical Analysis of Teacher Turnover

It is somewhat surprising that the DIDID-IV regression results in Section IV indicate that black students are not associated with higher turnover, so I present descriptive statistics about teacher turnover to put these regression results in perspective. Figure 5*a* shows the 1-year teacher turnover rates (leaving their current school) by year for those CM schools with BDs above and below the average. The first notable pattern is that although there are differences in turnover rates between low-BD and high-

[24] Appendix table B1 also shows results using serial correlation adjusted value-added estimates, which are qualitatively similar. Using the second lag of test scores to correct for measurement error in lagged test scores reduces the sample of teachers with estimated effects to less than half of those when one uses the lagged test scores as is. This would explain the additional noise.

[25] This calculation is based on estimates from Jackson and Bruegmann (forthcoming), who find that the coefficients on estimated standard normalized value-added estimates are 0.126 and 0.055 for math and reading, respectively, using these same data. Other studies have found that a one standard deviation increase in teacher quality increases student achievement by between 0.25 and 0.1 of a standard deviation (Rockoff 2004; Hanushek, Kain, and Rivkin 2005; Jacob and Lefgren 2008). The back-of-the-envelope calculations assume that teachers would have been as effective in their new schools as they were in their previous schools. If some of the estimated teacher value added is due to unobserved student characteristics or to the match between certain students and certain teachers, then a teacher's value added in one school may not be predictive of her value added in another school. Irrespective of how good a predictor estimated value added is across schools, if a school loses teachers with the highest estimated value added, it implies that there is a real reduction in teacher quality at that school.

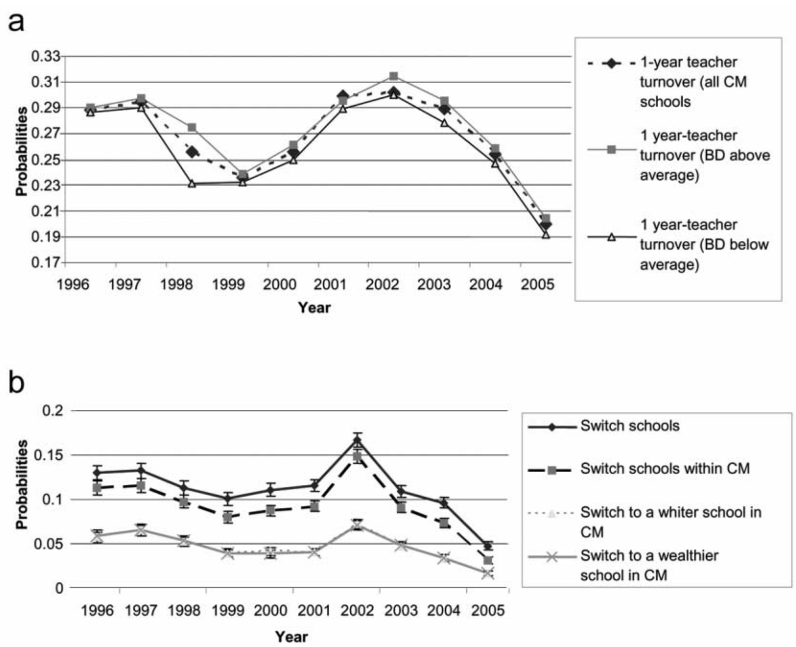

FIG. 5.—*a*, One-year teacher turnover rate in CM by BD (%). *b*, School-switching probabilities, all teachers in CM (+/− 2 standard errors).

BD schools (i.e., low-BD schools with low black enrollment shares have slightly higher turnover than high-BD schools with large black enrollments), the increases in turnover over time are almost identical for all schools. This is consistent with finding a statistically significant effect of percent black on turnover in the cross section but no statistically significant differential effect of percent black on turnover in the IV-DIDID estimates in columns 3 and 4 of table 3. The second notable pattern is that turnover is elevated for all CM schools between 2001 and 2003, suggesting that teachers may have been reacting to the change in student demographics and to the anticipated change in student demographics. Since a teacher sorting explanation would involve teachers switching schools rather than simply leaving their current school, figure 5*b* looks specifically at teachers switching schools. This panel shows a clear increase in teachers switching schools in 2002, which was obscured by looking at aggregate teacher turnover. Using simple *t*-tests, one can reject the hypothesis that switching was the same in 2002 as in 2001 or 2003 at the 5% level. The figure also shows that the vast majority of teacher switching was due to switching schools within CM rather than switching to schools outside the district. There is some evidence of increased switching to

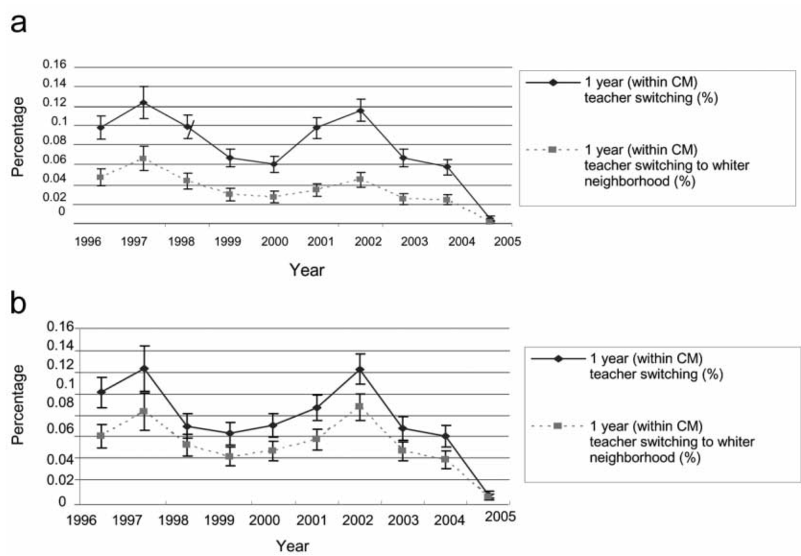

FIG. 6.—*a*, Within-CM switching rates for schools with BD above average (outflow of blacks; +/− 2 standard errors). *b*, Within-CM switching rates for schools with BD above average (inflow of blacks; +/− 2 standard errors).

whiter neighborhoods in 2002, but some of this may simply be due to mean reversion.

If teachers were switching schools in 2002 because they all preferred to teach in schools that had a lower share of black students, one would observe that (1) there was an increase in teacher turnover for those schools that had an increase in the share of black students, (2) there was a decrease in turnover for those schools that had a decrease in the share of black students, and (3) most of this change in turnover would be due to school switching. The dynamics documented in figure 6 show that this was not the case. Both high-BD and low-BD schools experienced an increase in teachers switching out of their schools to other schools in the district in 2002. This dynamic is much more consistent with there being heterogeneity in teachers' preferences for students, suggesting that some teachers liked teaching in schools with high shares of low-income minority students whereas other teachers did not.[26] This would also explain why the aggregate regression results show no differential change in teacher turn-

[26] Anecdotal evidence and conversations with district officials suggest that many teachers avoid inner-city schools because they find the working conditions difficult, whereas other teachers seek them out because they want to make a difference to students who really need the help.

over across schools despite a clear change in the characteristics of teachers within schools over time.

Readers may wonder if the movement of students systematically created job openings at schools that had an outflow of black students due to the policy change, leading to a change in teacher demand. An IV regression of the share of teachers that are new hires yields a coefficient on percent black of 0.009 and a standard error of 0.012. The standard error of the same OLS regression is 0.01, demonstrating that this lack of significance is not due to increased noise from the IV procedure. If teachers had no preferences for student demographics—since there were not disproportionately more new hires (i.e., vacancies filled) at schools that lost or gained black students as a result of the policy change—they would be no more likely than before to apply for transfers or leave their schools. Since the school district did not compel teachers to leave schools, the aggregate increase in teacher switching for all schools further suggests that the changes in mobility were likely due to a labor supply rather than a demand response.

VI. The Effect of the Policy Change on Incumbent Teachers and New Hires

The changes in the aggregate documented in Section IV may have occurred for three reasons: (1) schools that had an inflow of black students may have experienced an increased outflow of highly qualified teachers, (2) schools that had an inflow of black students may have found it more difficult to attract new highly qualified teachers than before the inflow of black students, or (3) some combination of the two. I attempt to disentangle these two margins by looking at changes in the characteristics of teachers who remained in a school (i.e., teachers who did not leave their school the previous year) and changes in the characteristics of newly hired teachers. All the analyses in this section use the IV-DIDID specification to remove potential endogeneity bias.

Table 4 reports the coefficient on percent black on the characteristics of individual teachers. Columns 1–9 are based on the sample of teachers who remained in their school from the previous year, and columns 10–18 are based on the sample of new teachers. Table 4 reports the IV-DIDID results. All models include year-by-district fixed effects, year-by-locale fixed effects, school effects, and post-by-BD effects. The results for incumbent teachers in columns 1–9 echo the aggregate teacher results. A school with a 10-point increase in the share of black students experienced a 1-year decline in the average years of experience among teachers who stayed in the school. Those teachers who stayed after the policy change were about 3 percentage points more likely to be black and had about 0.14 standard deviations lower value added in math and reading. These

Table 4
Effect of Changes in the Percentage of Black Students on Characteristics of Incumbent Teachers and New Hires

	Incumbent Teachers								
	Years of Experience (1)	Fewer than 4 Years' Experience (2)	4–10 Years' Experience (3)	11–20 Years' Experience (4)	More than 20 Years' Experience (5)	White (6)	Black (7)	Math Effect EB (8)	Reading Effect EB (9)
% black students	−.10729	.0037	.00226	−.00153	−.00349	−.00244	.00299	−.01368	−.0139
	[.03582]***	[.00155]**	[.00135]*	[.00109]	[.00154]**	[.00151]	[.00149]**	[.00415]***	[.00532]***
Observations	128,105	128,105	128,105	128,105	128,105	128,105	128,105	26,524	26,524
Number of schools	419	419	419	419	419	419	419	412	412
	New Hires								
	Years of Experience (10)	Fewer than 4 Years' Experience (11)	4–10 Years' Experience (12)	11–20 Years' Experience (13)	More than 20 Years' Experience (14)	White (15)	Black (16)	Math Effect EB (17)	Reading Effect EB (18)
% black students	.04085	−.00187	−.00058	.00105	.00061	−.00281	.00228	−.0007	−.0023
	[.06102]	[.00223]	[.00152]	[.00137]	[.00146]	[.00381]	[.00332]	[.01109]	[.01315]
Observations	24,464	24,464	24,464	24,464	24,464	23,969	23,969	2,580	2,580
Number of schools	419	419	419	419	419	419	419	345	345

Note.—Robust standard errors are in brackets. Standard errors are clustered at the zip code level. The sample is CM, Wake, Guilford, and Cumberland districts. All regressions are based on the same IV-DIDID specification detailed in eqq. (4) and (5). All regressions include year effects interacted with the school district, the decile of the school in the distribution of percent black in the neighborhood, and the locale. All specifications include school fixed effects and a POST × BD variable. The excluded instrument in these models is the POST × BD × CM variables interacted with the quintile of the school in the distribution of the percentage of black residents from the 2000 Census.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

results imply that within a school, those teachers who left schools that experienced an inflow of black students were on average more experienced and whiter and had higher value added than those who stayed.

Columns 10–18 look at the attributes of new teachers that schools hired. None of these point estimates are statistically distinguishable from zero, suggesting that either the sample of new teachers is too small to detect differences or there is no systematic difference in the characteristics of new teachers that schools hired after the policy change. However, the point estimates suggest that schools that experienced an inflow of black students were more likely to hire black teachers than they were before the policy change.

The results in table 4 suggest that white teachers, more experienced teachers, and teachers with high value added were more likely to leave schools that experienced an inflow of black students than black teachers, teachers with less experience, and teachers with low estimated value added. Direct tests for differential mobility across experience and value-added groups yield statistically insignificant results that are not generally robust across models. However, differential mobility by teacher race is a consistent finding across all models, and I present these results in table 5.

The dependent variable in table 5 is leaving the current school in the same year. The coefficient on percent black is reported, and all models include the full set of control variables in model (4) and instrument for percent black using 2SLS. Columns 1–3 show the effect for black teachers on leaving the current school and columns 4–6 show the results for white teachers. Columns 1 and 3, which use the three-way interaction as the excluded instrument, show that black teachers were 6 percentage points less likely to leave a school when the share of black students increased by 10 percentage points, whereas white teachers were 1.5 percentage points less likely to leave. The effect on black teachers is statistically significant at the 10% level, and the effect on white teachers (who were more numerous) is not significant at traditional levels. Columns 2 and 5 use the instrument interacted with the quintile of the percentage of black residents in the neighborhood as used in table 3. These results are largely the same, but now the effect for black teachers is statistically significant at the 5% level.

Since the analysis in Section V indicates that much teacher turnover and switching took place in 2002 rather than 2003 in anticipation of the change in student attributes, columns 3 and 6 use the percentage of black students the following year as the independent variable. The instruments are also altered so that POST_t denotes the year before students moved. The results from this model indicate that black teachers were about 1 percentage point less likely to leave a school when the share of black students was expected to increase by 10 percentage points; however, there is no statistically significant differential effect for white teachers.

Table 5
Difference in Mobility Response by Race: Dependent Variable: Leave Current School

	Black Teachers: IV-DIDID			White Teachers: IV-DIDID		
	(1)	(2)	(3)	(4)	(5)	(6)
% black students	−.00617	−.00761	. . .	−.00155	−.00159	. . .
	[.00325]*	[.00357]**		[.00181]	[.00175]	
% black students in the following year	. . .	. . .	−.00095	. . .	. . .	<.0001
			[.00043]**			[.00019]
Excluded instruments[a]	1	2	1	1	2	1
Observations	16,706	16,706	13,480	64,811	64,811	52,922
Number of schools	408	408	402	418	418	418

NOTE.—Robust standard errors are in brackets. Standard errors are clustered at the zip code level. Sample is CM, Wake, Guilford, and Cumberland schools. All regressions are based on the same IV-DIDID specification detailed in eqq. (4) and (5). All regressions include year effects interacted with the school district, the decile of the school in the distribution of the percentage of black residents in the neighborhood, and the locale. All specifications also include school fixed effects and a POST × BD variable.

[a] Instrument 1 is the three-way interaction BD × POST × CM, and instrument 2 is BD × POST × CM interacted with the quintile of the school in the distribution of the percentage of black residents in the neighborhood.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

Readers may wonder if all these differences were driven by sorting across race, that is, the movement of black teachers, who may have had, on average, weaker qualifications and lower value added. To determine if the changes in the aggregate were due to changes in the characteristics of white teachers, changes in the characteristics of black teachers, or simply changes in the racial makeup of teachers, I estimate equations (4) and (5) on mean teacher experience and mean teacher value added in math and reading for black and white teachers separately. Among white teachers, a 10-point increase in the percentage of black students is associated with a statistically significant 1.02 reduction in mean years of teaching experience and statistically insignificant 1% of a standard deviation decreases in mean value added for both math and reading. Surprisingly, among black teachers these effects are even stronger. A 10-point increase in the percentage of black students is associated with a statistically significant 1.74 reduction in mean years of teaching experience and statistically significant 30% and 39% of a standard deviation decreases in mean value added for math and reading, respectively. This implies that even though black teachers were more likely to stay in schools with growing black enrollment shares, those black teachers who left were better, on average, than those black teachers who remained. These results show that inflows of black students are associated with decreases in the average quality of both black and white teachers. This strongly suggests that sorting by student race occurs both across and within teacher race, implying that although teachers' preferences for student race may be associated with teacher race, there is substantial heterogeneity in teachers' preferences for student race among both white and black teachers.

In sum, schools that experienced an exogenous increase in the black enrollment share were relatively more likely to lose white teachers, experienced teachers, and effective math and reading teachers. The IV-DIDID estimates indicate that black teachers were less likely to leave schools, whereas white teachers were not differentially affected by an exogenous inflow of black students. Though the point estimates show that schools that had increasing black enrollment shares hired new teachers with lower estimated value added than before the inflow, these new hire results are not statistically significant.

VII. Concluding Remarks

The regression results show that the change from a race-based busing policy to a neighborhood-based controlled school choice model changed the student makeup of schools in Charlotte-Mecklenburg in a clear and foreseeable way. As predicted by the instrument, schools that had a greater share of black students than black residents in the surrounding neighborhood experienced an outflow of black students and an inflow of white students. The converse was also true. The sudden inflow or outflow of black students as a result of the policy change was associated with systematic changes in the makeup of teachers at the affected schools. Schools that experienced an increase in the black enrollment share saw a decrease in the proportion of experienced teachers, a decrease in the proportion of teachers with high scores on their licensure exams, and a decrease in teacher value added. I find that the aggregate decline in teacher quality in schools with increased black enrollment shares was due to these schools losing experienced and effective teachers. I find that white teachers were, on average, no more likely to leave schools that experienced an inflow of black students than to leave schools that did not, whereas black teachers were, on average, more likely to stay in schools that had an exogenous increase in the black enrollment share. This suggests that the relationship between teacher race and student race is not a mere artifact of colocation but likely the result of teacher preferences for student attributes that are correlated with race. While discrimination against black teachers at schools that had increasing white enrollment shares is consistent with the patterns for black teachers, it does not explain the fact that inflows of black students are associated with decreases in the average quality of both black and white teachers—making a discrimination explanation unlikely.

District employees assert that because there were no changes in CM hiring practices that accompanied the change in the student assignment policy, these changes in teacher characteristics were driven by teacher labor supply. In addition, empirical evidence supports this interpretation. Specifically, (1) new teacher hiring (vacancies) was not correlated with the direction of the flow of black students due to the policy change; (2) all schools in the district experienced increased turnover, suggesting a re-sorting of teachers rather than a general movement of teachers from certain schools to others with vacancies; and (3) teachers switched schools in anticipation of the demographic changes. Although I cannot definitively rule out a demand-side explanation, the bulk of the evidence supports a labor supply interpretation.

The dynamics of teacher turnover are consistent with a world in which some teachers prefer to teach in inner-city schools with low-income ethnic minority students and others prefer not to. These preferences appear to be correlated with teacher race, so that, on average, black teachers may have a greater preference for teaching in schools with larger shares of black students. However, I find that much sorting occurs within both white and black teacher populations. The theory of compensating differentials predicts that where teachers have heterogeneous preferences for student characteristics, if students are reshuffled (as they were), teachers would also re-sort across schools. In fact, this is exactly the type of dynamic one observes in the data. The fact that teachers who move and teachers who stay may have very different preferences suggests that estimates that look at changes in teacher behavior, on the margin, could grossly overstate or understate the overall or average effect of school characteristics on teacher mobility. This also suggests that compensating differentials estimated on the basis of mobile teachers may be very different from those for the average teacher.

Overall, the findings present some of the first compelling evidence that teacher characteristics and teacher quality are endogenous to student demographics. One can reject the hypothesis that the correlation between teacher quality and student demographics is merely an artifact of geography or residential segregation. The teacher sorting is probably responsible for some of the disparities in teacher qualifications that exist between low-income inner-city schools and affluent suburban schools, which in turn may be responsible for some of the cross-school achievement gaps that exist. The endogeneity of teacher quality with respect to student characteristics also suggests that the movement of effective teachers out of schools in predominantly black neighborhoods may be partially responsible for the increase in the black-white achievement gap associated with the end of school desegregation and residential segregation. An important implication of these findings is that policy makers should be cautious when advocating policies such as vouchers, school choice, district consolidation, or school busing that require the reshuffling of students across schools. Insofar as student characteristics affect where teachers teach, the change in teacher attributes caused by the reshuffling of students across schools needs to be taken into account when determining the overall anticipated effect of such policies.

Appendix A

Empirical Bayes Estimates

It has been pointed out that while teacher effects that come directly from (3) should yield consistent estimates of teacher value added under the identifying restrictions, these estimates are not the most efficient. A more efficient estimate of teacher value added is the empirical Bayes (EB) estimate that shrinks noisy value-added estimates toward the mean of the value-added distribution (in this case zero). Since the estimates are estimated with noise, then $\hat{\tau}_j = \tau_j + u_j$, where u_j is a random estimation error and $\tau_j \sim N(0, \mathrm{Var}(\tau))$, so that the total variance of the estimated effects is $\mathrm{Var}(\hat{\tau}_j) = \mathrm{Var}(\tau) + \mathrm{Var}(u_j)$. It is straightforward to show that in the presence of estimation error, $E[\tau_j | \hat{\tau}_j] = [\sigma_\tau^2/(\sigma_\tau^2 + \sigma_{u_j}^2)] \cdot \hat{\tau}_j$. The empirical analogue of this conditional expectation is an EB estimate of teacher value added.

I follow the procedure outlined in Kane and Staiger (2008) to compute the EB estimates. This approach accounts for the fact that (1) teachers with larger classes will tend to have more precise value-added estimates, and (2) there are classroom-level disturbances so that teachers with multiple classrooms will have more precise value-added estimates. To simplify the notation, I subsume all observed covariates into a single variable X_{jt} and drop the grade subscript g to rewrite equation (3) as

$$A_{ijt} - A_{ijt-1} = \eta X_{jt} + \tau_j + \upsilon_{jt} + \varepsilon_{ijt}. \tag{A1}$$

In (A1), the total error term is $z_{ijt} = \tau_j + \upsilon_{jt} + \varepsilon_{ijt}$. Since the student error component is equal to zero in expectation, the mean residual for classroom jt, $c_{jt} = \tau_j + \upsilon_{jt}$, contains the teacher effect and the idiosyncratic classroom error. Since the classroom error is randomly distributed, I use the covariance between the mean residuals of adjacent classrooms for the same teacher, $\mathrm{Cov}(c_{jt}, c_{jt-1}) = \hat{\sigma}^2_\tau$, as an estimate of the variance of true teacher quality. I use the variance of the classroom demeaned residuals as an estimate of $\hat{\sigma}^2_\varepsilon$. Since the variance of the residuals is equal to the sum of the variances of the true teacher effects, the classroom effects, and the student errors, I compute the variance of the classroom errors $\hat{\sigma}^2_\nu$ by subtracting $\hat{\sigma}^2_\varepsilon$ and $\hat{\sigma}^2_\tau$ from the total variance of the residuals.

For each teacher, I compute a weighted average of her mean classroom residuals, where classrooms with more students are more heavily weighted. Specifically, I compute

$$\bar{\tau}_j = \sum_{t=1}^{T_j} c_{jt} \cdot \frac{1/[\sigma^2_\nu + (\sigma^2_\varepsilon / N_{jt})]}{\sum_{t=1}^{T_j} \{1/[\sigma^2_\nu + (\sigma^2_\varepsilon / N_{jt})]\}}, \tag{A2}$$

where N_{jt} is the number of students in classroom jt, and T_j is the total number of classrooms for teacher j. To obtain an EB estimate for each teacher, I multiply the weighted average of classroom residuals $\bar{\tau}_j$ by an estimate of its reliability. Specifically, I compute

$$\hat{\tau}^{\mathrm{EB}}_j = \bar{\tau}_j \cdot \frac{\hat{\sigma}^2_\tau}{\hat{\sigma}^2_\tau + \sigma^2_{u_j}}, \tag{A3}$$

where $\sigma^2_{u_j} = (\sum_{t=1}^{T_j} \{1/[\sigma^2_\nu + (\sigma^2_\varepsilon / N_{jt})]\})^{-1}$ is the estimation variance of the raw value-added estimate. The shrinkage factor $\hat{\sigma}^2_\tau / (\hat{\sigma}^2_\tau + \sigma^2_{u_j})$ is the ratio of signal variance to total variance and is a measure of how reliable an estimate $\bar{\tau}_j$ is for τ_j.

This EB method, which has also been used in Rockoff (2004), Gordon, Kane, and Staiger (2006), and Jacob and Lefgren (2008), is intuitively appealing since it uses all the available information to "shrink" noisy teacher value-added estimates to yield efficient value-added estimates. While the results using the EB estimates are stronger than those using the raw value-added estimates, in practice, this adjustment does not qualitatively change the results.

Appendix B

Table B1
Regression Estimates of Test Score Growth

	Math (1)	Reading (2)		Math (1)	Reading (2)
Lagged score	−.2522	−.2594	Class size	−.0021	−.0011
	[.0036]***	[.0018]***		[.0004]***	[.0003]***
Peers: lagged score	−.1405	−.0873	School: urban fringe (large city)	.0042	.0476
	[.0077]***	[.0077]***		[.0202]	[.0178]***
Student: male	−.0085	−.0489	School: midsized city	−.042	.0005
	[.0015]***	[.0016]***		[.0205]**	[.0166]
Student: black	−.257	−.1879	School: urban fringe (midsized city)	−.0285	.0282
	[.0049]***	[.0043]***		[.0205]	[.0179]
Student: Hispanic	−.1077	−.0447	School: large town	.068	.0727
	[.0055]***	[.0055]***		[.0392]*	[.0337]**
Student: American Indian	−.2108	−.1294	School: small town	−.0157	.0381
	[.0070]***	[.0072]***		[.0223]	[.0191]**
Student: mixed	−.1563	−.0764	School: rural (inside CBSA)	−.0104	.0387
	[.0070]***	[.0074]***		[.0204]	[.0178]**
Student: white	−.1151	−.0396	School: rural (outside CBSA)	−.0074	.0344
	[.0045]***	[.0045]***		[.0197]	[.0174]**
Parental education: high school graduate	.1246	.1317	School: log enrollment	−.0051	.0041

Table B1 (*Continued*)

	Math (1)	Reading (2)		Math (1)	Reading (2)
	[.0023]***	[.0021]***		[.0099]	[.0078]
Parental education: some college	.1943	.2004	School: white (%)	.3944	.2903
	[.0033]***	[.0030]***		[.0891]***	[.0725]***
Parental education: professional graduate school	.2127	.223	School: Hispanic (%)	.3164	.226
	[.0032]***	[.0026]***		[.1261]*	[.1014]*
Parental education: junior college graduate	.2916	.291	School: black (%)	.1812	.1749
	[.0037]***	[.0027]***		[.0899]**	[.0733]**
Parental education: college	.3421	.3352	School: % free-lunch eligible	−.0093	−.0518
	[.0044]***	[.0032]***		[.0232]	[.0190]***
Parental education: graduate school	.3865	.3741	Year fixed effects?	Yes	Yes
	[.0062]***	[.0053]***			
Teacher and student are same race	.0022	−.0013	Grade fixed effects?	Yes	Yes
	[.0019]	[.0019]			
Teacher and student are same gender	.0058	−.0015	Observations	1,257,510	1,249,391
	[.0015]***	[.0016]			
Teacher: 0 years' experience	−.0478	−.0201	Number of teachers	30,974	30,888
	[.0196]**	[.0171]			
Teacher: 1–3 years' experience	−.0036	−.0063	Fraction of variance due to TFX	.321	.273
	[.0182]	[.0164]			
Teacher: 4–10 years' experience	.0118	−.0081	R^2	.18	.16
	[.0179]	[.0163]			
Teacher: 10–24 years' experience	.0177	.0008			
	[.0187]	[.0170]			
Teacher: 25+ years' experience	−.0015	−.007			
	[.0207]	[.0183]			

NOTE.—Robust standard errors are in brackets. All regressions include an indicator for missing parental education. The reference teacher experience group is teachers with missing experience data. Coefficients for the "other" student ethnicity category are suppressed. CBSA = core-based statistical area. TFX = teacher fixed effects.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table B2
Coefficient on the Percent of Black Students for Different Percentiles of the Value-Added Distribution

Percentile	Math	Math EB	Reading	Reading EB	Math Adjusted[a]	Reading Adjusted[a]
5	−.004 [.013]	.001 [.009]	−.014 [.013]	−.018 [.009]*	−.009 [.012]	−.021 [.011]*
10	−.015 [.009]*	−.008 [.008]	−.018 [.011]	−.022 [.009]**	−.011 [.010]	−.015 [.010]
15	−.013 [.010]	−.009 [.008]	−.014 [.013]	−.021 [.008]**	−.008 [.009]	−.009 [.009]
20	−.02 [.010]*	−.017 [.008]**	−.012 [.015]	−.02 [.009]**	−.009 [.009]	−.001 [.010]
25	−.014 [.010]	−.018 [.009]**	−.007 [.014]	−.019 [.010]*	−.01 [.009]	−.001 [.009]
30	−.014 [.008]*	−.016 [.007]**	−.011 [.012]	−.017 [.009]*	−.008 [.008]	−.002 [.008]
35	−.009 [.007]	−.021 [.008]***	−.012 [.008]	−.022 [.009]**	−.005 [.007]	−.002 [.007]
40	−.013 [.007]*	−.023 [.007]***	−.018 [.007]**	−.027 [.009]***	−.009 [.008]	−.005 [.006]
45	−.014 [.007]**	−.023 [.006]***	−.019 [.007]***	−.028 [.009]***	−.01 [.008]	−.008 [.006]
50	−.015 [.006]**	−.026 [.006]***	−.017 [.006]***	−.03 [.008]***	−.011 [.007]	−.006 [.005]
55	−.016 [.005]***	−.029 [.006]***	−.017 [.006]***	−.03 [.007]***	−.011 [.007]*	−.006 [.006]
60	−.013 [.005]**	−.028 [.006]***	−.015 [.006]**	−.029 [.006]***	−.011 [.006]*	−.005 [.006]
65	−.015 [.006]***	−.025 [.005]***	−.014 [.005]***	−.03 [.006]***	−.01 [.007]	−.006 [.007]
70	−.017 [.005]***	−.028 [.006]***	−.012 [.005]***	−.029 [.006]***	−.014 [.007]*	−.006 [.008]
75	−.016 [.007]**	−.027 [.005]***	−.016 [.006]**	−.028 [.007]***	−.018 [.008]**	−.011 [.009]
80	−.014 [.008]*	−.027 [.005]***	−.012 [.007]*	−.024 [.006]***	−.023 [.009]**	−.013 [.009]
85	−.015 [.009]*	−.023 [.005]***	−.013 [.008]	−.017 [.005]***	−.028 [.011]***	−.017 [.010]*
90	−.004 [.009]	−.018 [.005]***	.002 [.008]	−.013 [.005]**	−.016 [.011]	−.007 [.009]
95	−.009 [.010]	−.02 [.007]***	−.001 [.012]	−.013 [.007]**	−.012 [.012]	−.01 [.013]

NOTE.—Robust standard errors are in brackets. All regressions are based on the same IV-DIDID specification detailed in eqq. (4) and (5). All regressions include year effects interacted with the school district, the decile of the school in the distribution of the percentage of black residents in the neighborhood, and the locale. All specifications also include school fixed effects and a POST × BD variable. The excluded instrument is the three-way interaction BD × POST × CM interacted with the quintile of the school in the percentage of black residents distribution. Math and reading are the normalized value-added estimates that come directly from eq. (3). Math EB and reading EB are the empirical Bayes estimates from eq. (3).

[a] Math adjusted and reading adjusted are the normalized value-added estimates obtained from a 2SLS procedure that uses the second lag of test scores as an instrument for the first lag of test scores in eq. (3). The lack of statistical significance for these two outcomes reflects the fact that the sample of teachers with estimated value added under this method effectively shrinks by half.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics* 25:95–135.

Anthony, Emily, and Dan Goldhaber. 2007. Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *Review of Economics and Statistics* 89, no. 1:134–50.

Betts, Julian R., Kim S. Rueben, and Anne Danenberg. 2000. Equal resources, equal outcomes? The distribution of school resources and student achievement in California. Report, Public Policy Institute of California, San Francisco.

Bohrnstedt, George W., and Brian M. Stecher. 1999. Class size reduction in California: Early evaluation findings, 1996–1998. Year 1 evaluation report, CSR Research Consortium, American Institutes for Research, Palo Alto, CA.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management* 24, no. 1 (Winter): 113–32.

Brewer, Dominic J., and Ronald G. Ehrenberg. 1994. Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review* 13, no. 1 (March): 1–17.

Brewer, Dominic J., and Dan D. Goldhaber. 2000. Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis* 22, no. 2: 129–45.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review* 24, no. 4:377–92.

———. 2007. How and why do teacher credentials matter for student achievement? Working Paper no. 12828, National Bureau of Economic Research, Cambridge, MA.

Clotfelter, Charles, Helen F. Ladd, Jacob Vigdor, and Justin Wheeler. 2007. High poverty schools and the distribution of teachers and principals. *North Carolina Law Review* 85, no. 5 (June): 1345–80.

Gordon, Robert, Thomas Kane, and Douglas O. Staiger. 2006. Identifying effective teachers using performance on the job. White Paper no. 2006-01, Hamilton Project, Washington, DC.

Guryan, Jonathan. 2004. Desegregation and black dropout rates. *American Economic Review* 94, no. 4:919–43.

Hanushek, Eric A. 1997. Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis* 19, no. 2:141–64.

Hanushek, Eric A., John Kain, Daniel O'Brien, and Steven Rivkin. 2005. The market for teacher quality. Working Paper no. W11154, National Bureau of Economic Research, Cambridge, MA.

Hanushek, Eric A., John Kain, and Steven Rivkin. 2004a. New evidence about *Brown v. Board of Education*: The complex effects of school racial composition on achievement. Working paper, Hoover Institution, Stanford University.

———. 2004b. Why public schools lose teachers. *Journal of Human Resources* 39, no. 2 (Spring): 326–54.

———. 2005. Teachers, schools and academic achievement. *Econometrica* 73, no. 2:417–58.

Hanushek, Eric A., and Steven Rivkin. 2006. School quality and the black-white achievement gap. Working Paper no. 12651, National Bureau of Economic Research, Cambridge, MA.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2006. Preferences and heterogeneous treatment effects in a public school choice lottery. Working Paper no. 12145, National Bureau of Economic Research, Cambridge, MA.

Hastings, Justine S., Richard Van Weelden, and Jeffrey Weinstein. 2007. Preferences, information, and parental choice behavior in public school choice. Working Paper no. 12995, National Bureau of Economic Research, Cambridge, MA.

Hastings, Justine S., and Jeffrey Weinstein. 2007. No Child Left Behind: Estimating the impact on choices and student outcomes. Working Paper no. 13009, National Bureau of Economic Research, Cambridge, MA.

Hoxby, Caroline M. 2000. Peer effects in the classroom: Learning from gender and race variation. Working Paper no. 7867, National Bureau of Economic Research, Cambridge, MA.

Jackson, C. Kirabo, and Elias Bruegmann. Forthcoming. Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*.

Jacob, Brian A., and Lars Lefgren. 2008. Principals as agents: Subjective performance assessment in education. *Journal of Labor Economics* 26, no. 1:101–36.

Kane, Thomas J., and Douglas O. Staiger. 2008. Are teacher-level value-added estimates biased? An experimental validation of non-experimental estimates. Unpublished manuscript, Harvard University.

Kane, Thomas J., Douglas O. Staiger, and Stephanie Riegg. 2005. School quality, neighborhoods and housing prices: The impacts of school desegregation. Working Paper no. 11347, National Bureau of Economic Research, Cambridge, MA.

Lankford, Hamilton. 1999. A descriptive analysis of the New York State and New York City teaching force. Report prepared for the New York Supreme Court case *Campaign for Fiscal Equity v. New York State*.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis* 24, no. 1:37–62.

Lutz, Byron F. 2005. Post Brown vs. the Board of Education: The effects of the end of court-ordered desegregation. Finance and Economics Discussion Series Working Paper no. 2005-64, Federal Reserve Board, Washington, DC.

Rockoff, Jonah E. 2004. The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review Papers and Proceedings* 94, no. 2:247–52.

Scafidi, Benjamin, David Sjoquist, and Todd Stinebrickner. 2007. Race, poverty, and teacher mobility. *Economics of Education Review* 26, no. 2:145–59.

Todd, Petra E., and Kenneth I. Wolpin. 2003. On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113 (February): F3–F33.