TAKING RACE OUT OF THE EQUATION:

SCHOOL REASSIGNMENT AND THE STRUCTURE OF PEER EFFECTS

Caroline M. Hoxby

Gretchen Weingarth*

In the last and current decade, the Wake County school district reassigned numerous students to schools, moving up to five percent of the enrolled population in any given year. Before 2000, the explicit goal was balancing schools' racial composition; after 2000, it was balancing schools' income composition. Throughout, finding space for the area's rapidly expanding student population was the most important concern. The reassignments generate a very large number of natural experiments in which students experience new peers in the classroom. Using panel data on students before and after they experience policy-induced changes in peers, we explore which models of peer effects explain the data. We also review common models and econometric identification of peer effects. Our results reject the popular linear-in-means and single-crossing models as stand-alone models of peer effects. We find support for the Boutique and Focus models of peer effects, as well as for a monotonicity property by which a higher achieving peer is better for a student's own achievement all else equal. Our results indicate that, when we properly account for the effects of peers' achievement, peers' race, ethnicity, income, and parental education have no or at most very slight effects. We compute that switching from race-based to income-based desegregation has at most very slight effects, so that Wake County's numerous reassignments mainly affected achievement through the redistribution of lower and higher-achieving peers.

* This paper has its origins in Gretchen Weingarth's Harvard University senior honor thesis (2005), cited within. As early as 2003, Gretchen Weingarth recognized the useful variation generated by Wake County's reassignments. The thesis and this paper share the first half of the title, "Taking Race Out of hte Equation," which is quotation from Silberman (2003), cited within. Otherwise, the thesis and this paper are quite distinct. The authors are very grateful to the North Carolina Education Research Data Center for provision of and a great deal of help with data. The corresponding author is Caroline Hoxby, Department of Economics, Harvard University, Cambridge, Massachusetts 02138.

I. Peer Experiments in Wake County, North Carolina

Starting in the 2000-01 school year, the Wake County public school district switched from a desegregation plan that attempted to balance its schools on the basis of race to a plan that attempted to balance schools on the basis of family income. (Family income was measured by the percentage of students participating in the free or reduced-price lunch program.) Many students were reassigned to schools as a result. This was not Wake County's first venture into reassignment, however. The district had actively engaged in reassignment throughout the 1990s largely because growth in the county's population meant that previous assignment plans were continually outdated. Throughout the 1990s and the current decade, up to five percent of Wake County students were reassigned *in any given year*. The reassignments changed the peer composition of many school cohorts (a cohort is the group of students who are enrolled in the same grade in the same school in the same school year) and consequently the peer composition of many classrooms. We can identify changes caused by the reassignments, as opposed to potentially endogenous variation caused by a family's relocation, a student's switching to a private school, and similar phenomena. Before and after each classroom change, we observe the achievement of each student regardless of whether he was himself reassigned or stayed put and experienced a change because others were reassigned. In short, Wake County generates thousands of reassignment "experiments" and a unique opportunity to learn about how a students' achievement is affected by the peer composition of his class.

Our primary goal is to learn much more about the structure of peer effects work than has been learned previously. As a rule, it is a difficult empirical challenge to credibly identify the mere *existence* of peer effects, and–in consequence–most researchers have focused their attention on establishing existence. They typically use highly restrictive econometric specifications, most especially the linear-in-means model.[1] (Most researchers are well aware of the benefit of identifying the structure of peer effects

[1] See Manski (1993), Hoxby (2000), Brock and Durlauf (2001), and Graham (2004) for more discussion of the difficulties researchers typically encounter in the econometric identification of peer effects. The

but simply find it difficult to do with the data available.) The linear-in-means model assumes that each student has the same effect on each other student (a homogeneous treatment effect). It also assumes that a single student whose achievement raises a class's mean achievement by two points has precisely the same effect as several students whose combined achievement raises the class's mean by two points (that is, all effects operate through one moment: the mean of peers). The focus on establishing existence and the linear-in-means model in particular have been problematic because neither educational policy-makers nor economists would care much about peer effects if they merely existed and were linear in means. If peer effects were linear in means, then regardless of how peers were arranged, society would have the same average level of outcomes. Moreover, most applications of peer effects–school desegregation, school choice, college choice, urban economics–need to have non-linear peer effects to generate results that are interesting and that mimic the facts. For instance, several existing models generate stratification (segregation along the lines of ability) by assuming that peer effects exhibit single crossing–that is, a high achieving peer has more effect on another high achieving peer than she has on a low achieving peer. Other models assume that the peers who matter most are "bad apples" whose behavior disrupts everyone and triggers disruptive behavior from children who would otherwise be attentive.[2] The *structure* of peer effects matters greatly.

Classrooms are a good environment for the identification of this structure because (a) outcomes are reasonably well-defined, (b) students' incoming achievement and other characteristics (race, ethnicity, sex, poverty, native language, and disability) are recorded, and (c) classmates are actually forced to spend a large amount of time together. The last feature of classrooms is important because it means that if we

---

literature on peer effects is too dense for a thorough listing of good, recent empirical studies, but a partial list includes Sacerdote (2001), Stinebrickner and Stinebrickner (2001), Boozer and Cacciola (2001), Angrist and Lang (2002), Zimmerman (2003), Kremer and Levy (2003), Winston and Zimmerman (2004), Betts and Zau (2004), Carrell, Malmstrom, and West (2005). The papers most closely related to this one are Samms (2004) and Vigdor and Nechyba (2004).

[2] For a variety of different specifications (or implicit specifications) of peer effects, see Epple, Sieg, and Romano (2003), Kremer (1993), Nechyba (1996), Benabou (1996), Lazear (2001).

can locate exogenous variation in classroom composition, we have an experiment that can plausibly show the non-existence of peer effects. If we find that a student is unaffected when forced to spend 6 hours a day, 180 days a year in the company of another student, we can confidently assert that the latter student has had a small or no peer effect. In contrast, if we find that a person is unaffected when we merely put another person in his general vicinity (as might occur in a neighborhood, college, or workplace), we are not sure whether peer effects are weak or whether the two simply had no occasion to interact. Much of this paper is dedicated to our narrowing in on the structure of peer effects that best explains the data. We are much aided by the fact that Wake County reassignments changed the composition of classrooms in a wide variety of ways.

Our second goal is discovering whether desegregation on the basis of family income has different effects than racial desegregation. That is, we wish to evaluate Wake County's policy, albeit through the indirect method of carefully identifying race-based and income-based peer effects and then aggregating them. From the *Brown versus Board of Education* decision onwards, desegregation plans have been based on one of two arguments.[3] The first is that, regardless of claims about "separate but equal" schools, no district will provide truly equal resources to schools so long as they are segregated. In essence, this is an argument that policy makers will be willing to deprive schools that serve minority students to enrich schools that serve non-minority students so long as no non-minority students attend the former type of school and thus experience the deprivation. This argument depends on highly indirect peer effects that operate over a long period of time. The second argument is that the presence of non-minority students in the classroom has a direct, salutary effect on minority students. For instance, in the original *Brown* decision, the Court placed significant weight on evidence that, when segregated, minority students became depressed about their future prospects and therefore failed to achieve. As it turns out, the evidence on which the Court relied was slight and not credible by modern social scientific standards.

[3] See Ogletree (2004) for a review of the basis of the *Brown versus Board of Education* decision.

This is not to say that the second argument is wrong, but simply to say that it is a theory based mainly on anecdote and personal introspection. Indeed, desegregation policies in the United States, have *generally* been based on theories rather than evidence about peer effects. In this paper, we estimate achievement-based, race-based, and income-based peer effects simultaneously. We thereby determine whether there are effects that are truly raced-based or income-based, as opposed to race and income being mere correlates of achievement. We are able to answers questions such as, "Would we expect student X to have higher or lower achievement if we removed one non-poor black student from his class and inserted a poor White student whose achievement was equal to that of the student who was removed?" We also investigate whether peer effects run on dimensions such students' sex and parents' education.

In the next section, we briefly review popular models of peer effects. We describe Wake County's policy change and our data in Section III. In Section IV, we narrow in on the specification of achievement-based peer effects that best explains the data. This allows us to decide what models of peer effects are supported by the evidence. We investigate peer effects based on race, income, and other student characteristics in Section V . We also discuss the implications of our results for Wake County's policy change. In the final section, we conclude.

II. How Might Peer Effects Work?

The linear-in-means model proposes that a student's outcome is a linear function of the mean of his peers' outcome. The main appeal of the model is convenience: it can be estimated even when the amount of variation in the data is barely sufficient or when data are available only at an aggregate level. In addition, researchers appear to like the fact that the linear-in-means model treats all achievement symmetrically–almost as though the linear-in-means model were agnostic about how peer effects work. Agnostic, however, is what the linear-in-means model is not: it actually imposes strict assumptions about the forms peer effects take. Moreover, the linear-in-means model has the unappealing property that, if it were the true model, no form of segregation would be stable because all allocations of peers are equally beneficial in aggregate. Since certain forms of segregation crop up routinely, they are either due to institutional factors (having them play such a large role is unappealing) or due to a form of peer effects other than the linear-in-means model.

A formalization of the linear-in-means model that is appropriate for our application is:

$$y_{ijgt} = \beta_1 \bar{y}_{-i,jg,t-1} + \beta_2 \bar{X}_{-i,jg} + I_i^{student} \beta_3 + I_g^{grade} \cdot I_t^{school\ year}\ \beta_4 + \epsilon_{ijgt} \tag{1}$$

where y_{ijgt} is the outcome of student i in classroom j in grade g in school year t; $\bar{y}_{-i,jg,t-1}$ denotes the mean of his classmates' initial outcomes (from the end of the previous period); and $\bar{X}_{-i,jg}$ denotes the mean of his classmates' characteristics (such as race, gender, and income). The $\bar{X}_{-i,jg}$ characteristics are here assumed to be fixed over time, but could be time-varying without loss of generality. Note that the means on the right-hand side of the equation exclude the student himself–thus the "$-i$". The equation includes a full set of individual student fixed effects and a full set of grade-by-school year fixed effects. We discuss estimation of the model below.

Most other models of peer effects are defined on the basis of behavior, as opposed to the specification of an equation. Some popular ones are as follows.

The Bad Apple model of peer effects suggests that a presence of a single student with poor outcomes spoils the outcomes of many other students. If we find that an increase in the number of bottom-achieving students has a disproportionate negative effect on the achievement of students throughout the entirety of the distribution, we shall view this as evidence for the Bad Apple model. (By "disproportionate," we mean an effect substantially larger than the linear-in-means model would suggest.)

The Shining Light model of peer effects is the opposite of the Bad Apple model. It suggests that a single student with sterling outcomes can inspire all others to raise their achievement. If we see that an increase in the number of top-achievers has a disproportionate positive effect on the achievement of all other students, we shall take it as support of the Shining Light model.

A model implicit in some recent behavioral work is the Invidious Comparison model. In it, the advent of a higher achieving peer depresses the performance of everyone who is pushed to a lower rank in the local distribution (presumably by depressing their self-esteem). The advent of a lower achieving peer has the opposite effect: boosting the performance of all those who are pushed to a higher local rank.

The Boutique model of peer effects suggests that a student will have higher achievement whenever she is surrounded by peer with similar characteristics. This is essentially a model in which students do best when the environment is made to cater to their type. For instance, in schools, the Boutique model might mean that teachers organize lessons and materials around the learning style of a student if there is a critical mass of his type.

The Focus model of peer effects is closely related to the Boutique model but suggests that peer homogeneity is good for a student's learning, even if the student himself is not part of the group of homogeneous students. In this model, diversity is inherently disabling, perhaps because tasks cannot be well targeted to all students' needs. A bimodal distribution of peers may especially disabling because it may generate "schizophrenia" in the organization of work.

The opposite of the Focus model is the Rainbow model, so called because it suggests that all students are best off when forced to deal with all other types of students. The logic of the Rainbow model is that students learn the answer to a question more deeply when they see it approached from a variety of angles.

If we see that making a classroom more homogeneous is good for *all* students (even those who are consequently more anomalous), we shall take it as evidence for the Focus model. Naturally, we shall look upon opposite findings as evidence for the Rainbow model. If increased homogeneity only benefits students near the type that is becoming more prevalent, we shall take it as evidence of the boutique model.

We have already mentioned the Single Crossing model, which is probably less motivated by observed behavior than by the fact that it generates self-segregation in a mathematically elegant way. In the Single Crossing model, students with a higher initial level of the outcome are more sensitive to their peers' having a high level of the outcome. Thus, high achieving students benefit most and low achieving students benefit least from the presence of additional high achieving students. We can differentiate between Single Crossing and the Boutique model because, in the former, low achieving students benefit hardly at all from the presence of other low achieving students whereas, in the latter, they benefit substantially.

The Subculture model is, in some ways, the logical opposite of the Boutique model but it is likely only to affect certain minorities (achievement minorities, racial minorities, and so on). In the Subculture model, the majority type remains supportive of a minority person, such as a high achieving student or a black student, so long as he is relatively isolated. When, however, minority students become prevalent enough to form a critical mass, the majority type rejects them–perhaps because minority sub-culture threatens the environment that works best for the majority. The rejection could also be more benign. The majority may be willing to make sufficient effort to include a few minority members but unwilling to make the effort to include numerous minority members and also unwilling to include some minority students while rejecting others.

III. The Econometric Identification of Peer Effects in Wake County

To see, even in advance of the policy details, how Wake Country's experiments will help us identify peer effects, consider the linear-in-means model. The problems–self-selection, reflection, and measurement error/omitted variables–that plague it also plague other models.[4] The core of the model is:

$$y_{ijgt} = \beta_1 \bar{y}_{-i,jg,t-1} + \beta_2 \bar{X}_{-i,jg} + I_i^{student} \beta_3 + \ldots + \epsilon_{ijgt}. \qquad (2)$$

We have not written out the grade-by-school year fixed effects because they are not interesting: they are included mainly to eliminate nuisance variation in measured outcomes. Test scoring varies somewhat

[4] Manski (1993) coined the term "reflection problem," but it also known as the "the social multiplier" (Glaeser, Sacerdote, and Scheinkman, 2003).

from grade to grade and from year to year, and they soak up the resulting, uninformative variation.

Self-selection is the problem that a student who is going to have outcome $y_{ijgt}=y^*$ may seek out or be assigned to certain companions *because* of their initial outcomes. Their initial outcomes will then appear to cause the outcome y^* when the causality is actually the other way around. An obvious example is students in a cohort being divided into classes based on their initial achievement: high initial scorers in one class and low initial scorers in another. (A cohort is a grade-by-school year-by-school cell, whereas a class is a grade-by-school year-by-classroom cell. Most but not all cohorts have multiple classes.)

Reflection is the problem that occurs because, if peers influence a student, he also influences them. Thus, a student's own behavior is embodied in the outcomes of his peers. Because the mean $\bar{y}_{-i,jg,t-1}$ deliberately excludes student *i*'s own outcome, the equation already eliminates the purely mechanical incorporation of a student's own outcome into the mean. Nevertheless, the student's own outcome will make its way into the mean through his peers' outcomes. This is simply because each of them has an equation parallel to his own, with last period's test score being a function of his (previous) test score.[5] A concrete example is a mischievous student who induces other students to participate in his mischief. Even if he is the sole initial instigator (the child without whom no mischief would ever have occurred), he will have produced a crop of mischievous peers after a few grades. It would be hard for an outside observer to identify him as the instigator because he will appear to be part of a rascally gang.

The measurement error or omitted variables problem occurs because a determinant of the student's outcome is either measured poorly or omitted altogether, thus constituting part of ϵ_{ijgt}. If peers' characteristics are correlated with the measurement error or omitted variable, they will appear to cause the student's outcomes when they are really just proxying for his own characteristics. For instance, a

---

[5] If linear-in-means model holds, one can solve for the multiplier generated by the reflection problem. Unfortunately, each model of peer effects implies a different multiplier. Thus, unless one is interested in the linear-in-means model *per se*, there is little point in computing the multiplier associated with each estimated coefficient.

student's being poor is measured imperfectly by his participation in the free lunch program. If poor families tend to live together, then a child is quite likely to be poor himself if he does not participate in the lunch program but does attend school with many children who participate.

Equation (2) includes student fixed effects, and these are crucial because the set of identification strategies that are credible *conditional* on student fixed effects is quite different from the set of strategies that are credible without student fixed effects. With student fixed effects, we need only find variation in a student's classroom that is plausibly orthogonal to *time-varying* determinants of a student's achievement. The variation need not be orthogonal to time-constant determinants of the student's achievement, even if we cannot measure them. Consider: we compare a student before and after the composition of his class changes (the "treatment" changes). For our purposes, it is fine if the probability of experiencing a change in treatment is a function of the student's initial achievement and other fixed characteristics. What is *not* fine is if, conditional on this probability, the event of experiencing a change in treatment is related to time-varying factors that will affect his future achievement. (Readers familiar with the logic behind the propensity score will recognize this reasoning. We can condition on the probability of selection into treatment, so what we need for identification is that the event of treatment is random conditional on the probability of selection.) We shall argue that Wake County's experiments were of this type: while the probability of being reassigned or experiencing a reassigned peer was not random, it was based on relatively fixed student characteristics such as race and income. For a given set of fixed characteristics (for a given probability of being reassigned), we shall show that the actual reassignment event was apparently arbitrarily distributed.

For now, let us suppose that the argument is correct. How can we estimate equation (1) consistently? First, note that the student fixed effect absorbs all of student *i*'s time-constant determinants of achievement. Thus, we need not worry about such determinants being mismeasured or omitted.

Second, consider the formation of simulated instrumental variables for $\bar{X}_{-i,jg}$. If the policy-generated variation is as argued, then we want the simulated instruments to reflect reassignment-driven changes in the peer composition of a student's class. However, the simulated instruments must not reflect potentially endogenous student moves (such as occur when a family changes its residence or enrolls a child in private school). The simulated instruments must also not reflect assignment to classes within the cohort since such assignment (usually done by principals but influenced by teachers and parents) may be non-random. Define a student's "simulated instrument cohort" to be the group of students who would be in his cohort if reassignments are allowed but all potentially endogenous student movement is disallowed. Compute means based on the simulated instrument cohort, and use the resulting variables, $\bar{X}_{-i,sg}^{SimCo}$, as instruments for means based on a student's actual class. Note that s indexes the simulated cohort and displaces j, which indexes classrooms. The superscript "*SimCo*" is just a forcible reminder that the cohort is the simulated, not actual, one.

If a student's cohort does not experience policy-based reassignments, such instruments will be constant over time, will be soaked up by his individual fixed effect, and will contribute nothing to the estimates. This is appropriate because the student has experienced no credibly exogenous variation in peers. Note that the instrument is at the (simulated) cohort level, not the class level. Thus, endogenous assignment within the cohort does not affect the estimates. (Formally, the intention to treat varies only at the simulated cohort level. Thus, the instrumental variables estimates is a treatment-on-the-treated effect that reflects the difference between treated and untreated compliers. Compliers are students who experience at least as much of the new type of peer as they did before the new arrival of the reassigned peers.)

Third, consider the formation of simulated instrumental variables for $\bar{y}_{-i,jg,t-1}$. We can proceed along similar lines to form the instruments except for the fact that outcomes change over time and these changes may embody the reflection problem.[6] We fix this problem simply by forming the instrument

---

[6] To see this, suppose that a student is randomly assigned to experience some new peers in his simulated instrument cohort. Over time, he influences their achievement and their altered outcomes would used next period to compute simulated instruments if we did not based the instruments on initial achievement.

based on the *initial* achievement of each peer. If there is no change in peers in the simulated instrument cohort, there is no variation in the instrument and it is soaked up by the student's fixed effect. If the simulated instrument cohort changes because of reassignment, the reflection problem does not occur because the reassigned peers had not experienced the student when their initial achievement was determined.

Summing up, our first and second stage equations for estimating the linear-in-means model are:

(3)
$$\bar{y}_{-i,jg,t-1} = \alpha_1 \bar{y}^{SimCo}_{-i,sg,t_0} + \bar{\boldsymbol{X}}^{\boldsymbol{SimCo}}_{\boldsymbol{-i,sg}} \boldsymbol{\alpha}_2 + \boldsymbol{I}^{\boldsymbol{student}}_{\boldsymbol{i}} \boldsymbol{\alpha}_3 + \boldsymbol{I}^{\boldsymbol{grade}}_{\boldsymbol{g}} \cdot \boldsymbol{I}^{\boldsymbol{school\ year}}_{\boldsymbol{t}} \boldsymbol{\alpha}_4 + \xi_{sgt} + \xi_{ijgt}$$

$$\bar{X}^{k}_{-i,jg} = \lambda^k_1 \bar{y}^{SimCo}_{-i,sg,t_0} + \bar{\boldsymbol{X}}^{\boldsymbol{SimCo}}_{\boldsymbol{-i,sg}} \boldsymbol{\lambda}^k_2 + \boldsymbol{I}^{\boldsymbol{student}}_{\boldsymbol{i}} \boldsymbol{\lambda}^k_3 + \boldsymbol{I}^{\boldsymbol{grade}}_{\boldsymbol{g}} \cdot \boldsymbol{I}^{\boldsymbol{school\ year}}_{\boldsymbol{t}} \boldsymbol{\lambda}^k_4 + \upsilon^k_{jgt} + \upsilon^k_{ijgt} \quad \forall \quad X^k \in \boldsymbol{X}$$

$$y_{ijgt} = \beta_1 \bar{y}_{-i,jg,t-1} + \bar{\boldsymbol{X}}_{\boldsymbol{-i,jg}} \boldsymbol{\beta}_2 + \boldsymbol{I}^{\boldsymbol{student}}_{\boldsymbol{i}} \boldsymbol{\beta}_3 + \boldsymbol{I}^{\boldsymbol{grade}}_{\boldsymbol{g}} \cdot \boldsymbol{I}^{\boldsymbol{school\ year}}_{\boldsymbol{t}} \boldsymbol{\beta}_4 + \epsilon_{sgt} + \epsilon_{ijgt},$$

where t_0 is the period in which we initially observe the student. Notice the double error terms in each equation. These remind us that the equations must be estimated with robust standard errors clustered at the level of simulated instrument cohort.

We have described our identification strategy using the linear-in-means model for completeness, but in fact we shall use a large number of moments other than means. The strategy is, however, precisely parallel for each of the moments in question.

IV. Wake County's Reassignment Plans

In 1954, the U.S. Supreme Court called for the end of racial segregation in schools in its *Brown versus Board of Education* decision. Integration efforts through the first half of the 1960's remained weak, in part due to resistance.[7] In the second half of the 1960s, however, a combination of forceful court decisions–for instance, *Green versus New Kent County, Virginia*–and financial incentives from the federal

[7] Most famously, in 1957, black students were assigned to a predominately white school in Little Rock, Arkansas. The National Guard was pressed into service to ensure the students' safety.

government caused school districts to begin reassignment, busing, and similar involuntary methods of balancing the racial composition of schools.[8]

Like many other Southern districts, Wake County began substantial efforts at racial desegregation in 1965, shortly after the Elementary Education Act made the district choose between receiving substantial new federal funds or staying segregated. Wake County implemented a race-based reassignment plan, the goal of which was that each school should reflect the racial composition of the county.[9] District administrators divided the county into geographic nodes (there are currently about 700, each with an average of 150 students). The children in each node all follow the same reassignment plan, if any. Thus, the characteristics of an *individual* student are never a factor in his being reassigned. Throughout the 1990s, Wake County selected nodes for reassignment to balance schools' racial composition and for other reasons described below. As many as 5,500 students, or 5 percent of the district's students, were reassigned in a single year.

In 1994, the U.S. Supreme court ruled that the practice of desegregating schools based solely on race fell outside of the Equal Protection clause of the Fourteenth Amendment. In *Shaw versus Hunt*, a 1996 ruling, the Court stated that race could not be the “dominant and controlling consideration” in making reassignment decisions.[10] Most dramatically, in 1999 (*Tuttle versus Arlington County School Board)*, the Courts disallowed school districts from considering race in their decision to assign students to

---

[8] See Reber (forthcoming) and Clotfelter (2004) for a review of the evidence on court-ordered desegregation. Interestingly, Cascio, Gordon, Lewis, and Reber (2005) show that much desegregation was *not* caused by court orders but rather by the federal government's tying Title I funds to desegregation efforts. A consequence of the Civil Rights Act and the Elementary Education Act of 1965 was that Southern schools stood to gain substantial funds if they complied with federal guidelines on desegregation.

[9] Specifically, Wake County schools were supposed have black shares between 15 and 45 percent, a range centered on the 30 percent black share in the county's schools when desegregation began. For much of the detail on Wake County's policy, we rely on Weingarth (2005) and Silberman (2002).

[10] Specifically, the Court ruled that legislature could be conscious of a student's race when making reassignment decisions, but that race could not be the "dominant and controlling consideration."

schools.

By the late 1990s, Wake County's Board of Education believed that they had to change their method of desegregation or risk legal challenges of their own. Starting with the 2000-01 school year, they switched to reassigning students on the basis of family income rather than race. The goal of the new plan was balancing the schools' percentages of students participating in free or reduced-price lunch program. The target was 40 percent, the percentage of Wake County's students who participated in the lunch programs in 1999-00. During the years of the new plan that are covered by our data, as many as 4,157 students were reassigned in a single year. In more recent years not covered by our data, even larger numbers have been reassigned: up to 11,000 in a single year.

A. Practical Reassignment

If Wake County had merely reassigned nodes to balance schools' racial composition (up through 1999-00) or income composition (from 2000-01 onwards), its task would have been quite simple. But, in practice, when policy makers consider a node, their decision about reassignment is greatly dictated by school over- and undercrowding, existing bus routes that could be modified to link overcrowded to undercrowded schools, construction projects that displace existing students, and the advent of new buildings. This is because the Raleigh-Durham metropolitan area grew rapidly in the 1990s and continues to grow rapidly today: Wake County alone experienced enrollment growth of 44,718 students–a sixty percent increase–between 1990 and 2003.[11] Aligning students with available space and reasonably efficient bus routes was crucial; balancing schools' race or income composition was desirable but not paramount. In a typical year, it appears that only about 16 percent of reassignments were based purely on balancing considerations.[12]

---

[11] Wake County Government, education statistics internet site. http://www.wakegov.com/county/planning/demographic/dd_Education.htm

[12] The Wake County Public Schools Office of Growth and Management lists the following factors, in order, as the basis for reassignment: the opening of new schools; crowding at existing schools; the expansion of year-round schools; construction on, improvements to, and expansion of existing school

Most families comply with reassignment partly because Wake County attempts to run *all* of its schools well and partly because noncompliance is difficult. Assignments are not announced until May 15 of each year. Parents then have a fortnight to submit a transfer request (an appeal of the assignment), knowing that–if a transfer is approved–they will thereafter have to provide transportation to the school themselves. The only transfer requests that have a high probability of success are those in which parents have picked an alternative school that is under-filled or whose balance is such that the arrival of their child will help the school reach its target.[13] Wake County makes it hard for a parent to predict what the reassignment plan will be and take strategic steps in advance (such as by moving). Node maps are not published, and data on the characteristics of nodes that Wake County uses in the assignment decision are not publically available. While anyone can look up the current year school assignment for any given address, parents cannot obtain a spreadsheet of addresses and assignments–even for the current year, let alone for a sufficient number of previous years to conduct a proper analysis. Each year's preliminary and final assignment plan is removed from the internet when the official comment period is over.

For our purposes, the bottom line is as follows. First, both before and after 2000-01, students with the same characteristics who attended schools with similar race and income composition characteristics might experience arbitrarily different treatments. Indeed, we show below that, once we condition on a student's race, ethnicity, lunch participation, and initial school and grade (all of which are absorbed by the fixed effect in our analysis), the event of being reassigned appears to be quite random. In

---

facilities; transportation and travel time and distance; the transportation required to attend a magnet school; diversity indicators; the percentage of students who qualify for free or reduced-price lunches; recent trends in enrollment growth; reading achievement of students.

[13] Wake County has a system of magnet schools to which students can apply, but all such applications are submitted and approved or disapproved *before* reassignments are announced on May 15. Thus, a family that does not like its assignment cannot apply to a magnet school for the upcoming school year. Moreover, the magnet schools are meant primarily to help the district achieve racial (pre-2000) or income (post-2000) balance, so a parent who wants to avoid his child being sent to a more balanced school cannot generally achieve this by applying to a magnet school. Students who select magnet schools are treated as potentially endogenous movers, as are students who self-select into year-round schools. Note that many students are simply assigned to year-round school.

particular, we find that reassignment is not a function of a student's initial score. Although arbitrary assignment to treatment (conditional on a student's fixed characteristics) may annoy Wake County parents, it is useful for econometric identification. Second, a node's treatment over time tended to shift even if the node itself remained the same. This was partly due to the change in Wake County's desegregation policy, partly due to school renovation and the addition of buildings, and partly due to how changes elsewhere in the county affected local bus routes. Because nodes' treatment changes over time, students who are untreated in one part of the sample (for instance, earlier years) experience treatment in another part. This helps to guarantee that the control group is helpful for estimating the counterfactual–in other words, the grade-by-year effects, school effects, and so on. Also, the changing treatment of nodes makes families more likely to comply with reassignment: moving to an untreated node is no guarantee of remaining untreated. Third, the majority of children (62.4 percent) in Wake County experienced a change in peer composition purely because of reassignment. Of these, 38 percent were themselves reassigned and 62 percent were part of a cohort affected by other students' reassignment.

B. Data

We use data on third through eighth graders in Wake County from the 1994-95 through 2002-03 school years. We are grateful to the North Carolina Education Research Data Center, whose staff graciously provided data they had carefully compiled.[14] Our primary measure of achievement is a student's score on North Carolina' statewide end-of-grade tests. The dataset includes measures of race, gender, free and reduced-price lunch participation, and (rather unusually for administrative data) parents' education.[15]

---

[14] Weingarth (2005) contains considerable detail on the dataset. It is also described in documents posted on the website of the North Carolina Education Research Data Center.

[15] Prior to 1998-99, North Carolina did not record participation in free or reduced-price lunch in its state database. This does not affect our analysis, for two reasons. First, during the period in the which the desegregation plan was based on lunch participation, we do have the measure. Second, to the extent that we need a measure for prior years, we backcast a student's lunch participation or predict it using his parents' education. Our backcast and predicted measure matches up well with school-level participation

We use a student's total (reading plus math) scale score. This is simply because the results for reading and math were very similar so that the total is an informative summary statistic. We ensure that the test scores are comparable over time for the purposes of analysis. We do this by, first, using the same official scale for all years, and, second, including grade times school year effects in all our estimations to pick up idiosyncratic changes in the test or scoring.[16] We identify a student's actual classroom peers by identifying the group who share the same teacher code in the same cohort (grade, school, and school year). According to the North Carolina Education Research Data Center, these codes properly identify classrooms with at least 95 percent accuracy. The small degree of inaccuracy does not concern us because it appears to be random measurement error and we are instrumenting for classroom composition with cohort composition anyway. There is negligible error in a child's recorded cohort.

We classify all year-to-year transitions for a student into: staying, being reassigned by policy, making a feeder school transition (this occurs when all students from a certain elementary school are automatically "fed" into a certain middle school), or moving for a potentially endogenous reason. Using these classifications, we construct an indicator of each student's simulated instrument cohort in each year. Remember that the simulated instrument cohort allows policy-based moves but disallows all potentially endogenous moves, thereby "keeping" movers with their prior cohort.[17]

Table 1 shows descriptive statistics for our data. Note especially that the total test score has a standard deviation of 24.4–this number will be useful for assessing the magnitude of our results.

Table 2 shows the results of a linear probability regression that demonstrates that, once we condition on a student's race, ethnicity, free or reduced-price lunch participation, and initial school,

---

data, which is available prior to 1998-99.

[16] North Carolina introduced new scales for math in 2000-01 and for reading in 2002-03. We use the published conversion table between the old and new scales to put all scores into the old scales.

[17] Of potentially endogenous movers, we observe 92 percent both before and after the move because they move within North Carolina. In any case, observing them before and after is not terribly important owing to the instrumental variables strategy.

experiencing a policy-driven change in one's peers is not statistically significantly correlated with prior achievement or parents' education. We present another linear probability regression, the estimates from which indicate that, once we condition, being reassigned is not statistically significantly correlated with prior achievement or parents' education. These findings suggest that the staff in charge of reassignment used the variables they were supposed to consider (race, lunch participation, geography) but did not discriminate among students along dimensions they were not supposed to consider.[18] As a result, it is reasonable to assume that treatment (experiencing a policy-driven change in one's cohort) was random conditional on a student's fixed characteristics.

V. Understanding the Structure of Peer Effects

In this section, we clarify which specifications embody peer effects well. We consider peers' achievement only. That is, all of our explanatory variables are moments based on peers' initial achievement. In the next section, once we have settled on a specification that fits the achievement data well, we shall add explanatory variables based on peers' race and ethnicity, lunch participation, and so on.

Before proceeding, it is useful to point out that all of the implied first stage regressions have very ample explanatory power. This should come as no surprise because the instruments are constructed to capture all of the variation in the peer variables *except* the variation caused by potentially endogenous moves. The coefficient of interest in each implied first stage regression (the coefficient on the simulated instrument corresponding to the dependent variable) is always estimated to be positive and is always highly statistically significant. The vast majority of such coefficients are about 0.8, although a few are as low 0.25. The vast majority of associated t-statistics are over 100, although a few are as low as 30.

---

[18] The Wake County rules allow the authorities to consider students' prior achievement when making reassignment decisions. Table 2 does not show, however, any evidence of such consideration. (Indeed, the point estimates suggest that, if anything, low achievers are less likely to be reassigned, which is the opposite of what people commonly expect.) We think it is likely that, at the node level, there is insufficient *persistent* variation in achievement (conditional on everything else listed) for the authorities to base their decisions upon it.

A. Specifications in which Peers have Homogeneous Treatment Effects

To facilitate comparisons with other research on peer effects, we estimate the linear-in-means model, both by least squares and simulated instrumental variables. Results are displayed in Table 3, which also shows two other specifications estimated by simulated instrumental variables. Although the other two specifications allow for a variety of peer effects that the linear-in-means model does not, all the specifications shown in Table 3 have one thing in common: they restrict peers to have homogeneous treatment effects. That is, each student affects all of his peers identically, regardless of how similar he is to them initially.

We show ordinary least square results in the left-hand column, purely for interest. They suggest that a student's score is unaffected by the mean of his class's previous year test scores, controlling for student and other fixed effects. We have already mentioned that least squares estimates are highly problematic, so we shall proceed without interpreting the estimate further.[19]

The simulated instrumental variables estimate of the linear-in-means model suggests that adding peers who raise mean achievement by one point raises a student's own achievement by about 0.25 points. This effect is statistically significant and demonstrates the utility of an empirical strategy that excludes endogenous variation. The estimated effect is also quite large, though well within the range of previous estimates. Given our earlier discussion, however, one hardly knows what to do with the number. It cannot be used as the undergirding for most models of choice and it is difficult to use it to evaluate Wake County's desegregation policies (since all policies produce the same aggregate outcomes in the linear-in-means world and there is no explicit social welfare function with which to value gains and losses among students).

Because it is plausible that low-achieving and high-achieving students do not affect others purely

[19] One problem we did not mention (because it is irrelevant when we use instrumental variables) is regression to the mean. Thus, if a class does poorly one year because of some shock, its members can be expected to do well the next year simply because they are returning to their true level of achievement. Such phenomena can cause least squares estimates, like the one shown, to be downward biased.

through their effect on the mean, we relax the linear-in-means specification to include three additional moments: the shares of classmates with initial test scores in the bottom quartile, second quartile, and top quartile of the *countywide* distribution.[20] (The share with initial scores in the third quartile is omitted for obvious reasons.) The mean and the three other moments are instrumented with variables based on the simulated instrument cohort.

The results are somewhat confusing. It still appears that higher initial mean scores among classmates raise a student's own score: a 1 point increase in the mean raises his own score by 0.35 points. Also, if the share of his class with score in the second quartile rises by 10 percent, his own score falls by 3.2 points (13 percent of a standard deviation). The latter result in particular seems too large, and it is also hard to reconcile with the remaining results: the share of classmates with scores in the bottom quartile has no effect and the share of classmates with scores in the top quartile has a negative and statistically significant effect. Specifically, if the share of a student's class with scores in the top quartile rises by 10 percent, his own score falls by 1.3 points (5 percent of a standard deviation). While it is possible to construct peer effect models that reconcile these odd results, one cannot do so with models in which treatment effects are homogeneous–a restriction we have so far imposed. For instance, the linear-in-means, bad apple, and shining light model are all clearly rejected. (We can reject the linear-in-means model formally. The χ^2_3 statistic on the test is 50.5 with a p-value less than 0.0000.) The evidence is also hard to reconcile with the Invidious Comparison model: while a greater share of very high achievers has the expected negative effect, a greater share of very low achievers has no effect. The findings are also incompatible with the Rainbow model because that model suggests that adding students at both ends of the distribution should raise everyone's performance.

To drive home the point, we estimate an even more augmented specification, the estimates from

---

[20] That is, we computed percentiles of the *countywide* distribution of test scores for each grade and school year. We compare students' scores to these percentiles. Thus, the moments in question indicate the percentage of students in the class who are low or high achievers by a standard that is fairly absolute (certainly not closely related to the class's or school's own performance).

which are shown in the right-hand column of Table 3. It includes, in addition to peers' initial mean test score, the shares of classmates with initial test scores in each decile of the countywide distribution. (The share with initial scores in the bottom decile is omitted, and we instrument for all the achievement variables.) We can discern no sensible pattern in the results. A larger share of peers in second, seventh, and eighth deciles apparently raises a student's performance, but the fourth decile has a (borderline significant) negative effect. The remaining coefficients are statistically insignificant, but even the relative magnitudes and signs of the set of point estimates are difficult to align with one or more peer effect models. We again soundly reject the linear-in-means model. The χ^2_9 statistic on the test is 84.2 with a p-value less than 0.0000. Overall, we conclude that the data provide little support for models in which peers have homogeneous treatment effects.

B. Specifications in which Peer Effects Depend on the Student's Own Achievement

We now turn to specifications in which we allow the effects of peers to vary with a student's own initial achievement. Specifically, we associate each student with his initial score's decile in the countywide distribution of scores. Indicators for each student's decile are fully interacted with the ten variables representing the shares of classmates with initial test scores in each decile of the countywide distribution. Formally, the equation we estimate is:

(4)

$$y_{ijgt} = \gamma_1 I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 1} + \gamma_2 I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 2} + \ldots + \gamma_{10} I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 10} + \gamma_{11} I_{i,jg,t-1}^{decile\ 2} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 1} + \ldots + \gamma_{100} I_{i,jg,t-1}^{decile\ 10} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 10} + \boldsymbol{I_i^{student}} \boldsymbol{\gamma_{101}} + \boldsymbol{I_g^{grade}} \cdot \boldsymbol{I_t^{school\ year}} \boldsymbol{\gamma_{102}} + \varepsilon_{sgt} + \varepsilon_{ijgt},$$

where $I_{ijg,t-1}^{decile\ 1}$ is an indicator for student *i*'s test previous year score being in the bottom decile of the countywide distribution and $I_{-i,jg,t-1}^{decile\ 1}$ is the mean of the same indicator for his classmates. Keep in mind that all of the achievement-based explanatory variables are instrumented. For instance, the simulated instrumental variable constructed for $\bar{I}_{-i,jg,t-1}^{decile}$ is $\bar{I}_{sg,t_0}^{decile\ 1,\,SimCo}$.

Equation (4) is a very flexible functional form that can do a good job of representing the Invidious Comparison model, the Boutique model, and the Single Crossing model. The specification cannot, however, represent the Focus, Rainbow or Subculture model well because each of these models posits that the effect of a peer on another student is not merely a function of their (the twosome's) characteristics, but also a function of the achievement distribution in the remainder of the class.

It is not elucidating to present one hundred coefficients in a table, so we plot them. Figure 1 shows them all, and Figure 2 is a close-up of sorts. Note that the coefficients are identified only up to a constant so that, while the units on the vertical axis are meaningful, the position of each line relative to zero is not. Readers should concentrate on the shape of each line as it proceeds from the left- to the right-side of the figure.

In Figure 1, the coefficients plotted on the white background (toward the middle of the figure) tend to be statistically significantly different from zero at the 0.2 level at least. As we move out from the center of the figure, standard errors tend to grow. This occurs because there are many "experiments" in which a class receives a substantial boost in its share of peers who are middling, but few experiments in which a class receives a substantial boost in its share of peers who are very bottom or very top performers. Nature does not distribute very bottom and very top performers in such a way that they arise in neat clusters associated with geographic nodes. In the lightly shaded areas on either side, the estimates are so noisy that they should be taken with a very generous pinch of salt: their standard errors are typically three-quarters of the absolute value of the point estimate. In the deeply shaded regions on the outside of the figure, the estimates are very noisy: their standard errors are as much as 2 times the absolute value of point estimate. We show the estimates in the deeply shaded areas for completeness only: readers should avoid anything resembling literal interpretation. We do not show the deeply shaded regions at all in Figure 2 or the subsequent figures.

Figure 1 includes a great many estimates, but some patterns are immediately discernible. Consider the line based on students who are themselves initially in the bottom decile. Ignoring the estimate in the deeply shaded regions, we see that bottom decile students receive the greatest benefit when reassignment boosts the share of classmates in the second and third deciles. A ten percentage point increase in the share of peers who score at the 15^{th} percentile generates 4.5 more points on the test than the same size increase in the share of peers who score at the 85^{th} decile. 4.5 points is 18.5 percent of a standard deviation. At the other end of the spectrum, students who themselves are initially in the top decile benefit most when reassignment boosts the share of classmates in the fifth through ninth deciles. A ten percentage point increase in the share of peers who score at the 85^{th} percentile generates ten more points on the test than does the same size increase in the share of peers who score at the 15^{th} decile. Ten points is 40 percent of a standard deviation. Students who fall between the two ends of the spectrum have lines that lie between the two extreme lines just described. It is easier to see effects if we eliminate some lines, and this is what we do in Figure 2.

It shows that students who themselves initially score in the ninth decile exhibit much the same pattern as students who initially score in the top decile: increases in the shares of high performing peers are most beneficial. The line, however, for students in the seventh decile (61^{st} through 70^{th} percentiles) is much flatter. While it probably does peak at the 75^{th} percentile, suggesting that raising the share of such peers is most helpful, the difference between boosting peers at the 75^{th} and 25^{th} percentiles is negligible. Similarly, for students in the third decile, boosting peers at the 25^{th} percentile appears to be most helpful, but boosting peers at the 65^{th} or 75^{th} percentile is not much worse. Oddly, all of the lines for "interior" students exhibit a mild U-shape, suggesting that boosting the share of classmates who score near the 50^{th} percentile is least beneficial. This pattern is hard to understand, especially for the fifth decile students who themselves score in this range.

On the whole, we see substantial support for the Boutique model in Figures 1 and 2. Students who themselves exhibit the extremes of initial achievement benefit from the (net) arrival of like scorers. We see little evidence for either the Shining Light or the Bad Apple model: a boost in the share of very high or very low scorers has a mixed effect, not a uniform and disproportionate effect. We also see little evidence for the Invidious Comparison model: consider the lines for students who initially score in the third, fifth, and seventh deciles. They seem to benefit (mildly) from boosts in the share of both lower and higher scorers. The Single Crossing model also gets little support. While students who are themselves initially high scoring do seem especially sensitive to the performance of their peers, so do students who are themselves initially low scoring. Insensitivity to peers is apparently most characteristic of students who are themselves initially middling.

Figures 1 and 2 do not help us evaluate the remaining models because they require a specification that allows each of the coefficients described above to vary with the distribution of achievement in the rest of the class. Such a specification tests the limits of our data: every relaxation of the functional form cuts the number of "experiments" identifying each coefficient. Nevertheless, some additional relaxation seems warranted because puzzles remain. The Boutique model, for instance, cannot explain the mild U-shape described above: the Boutique model suggests that students who are themselves near the 50th percentile should benefit especially from increases in the share of classmates in the middle deciles.

C. Specifications in which Peer Effects Depend on a Student's Own Achievement and the Distribution of Achievement in the Rest of His Class

We estimate a augmented version of the previous equation in which peer effects may differ among classrooms of three types: those whose initial median score is low (in the bottom third of classroom medians, around the 25th percentile of the student population score), medium (in the middle third of classroom medians), and high (in the top third of classroom medians, around the 75th percentile of the student population score). This gives us the flexible equation:

(5)
$$y_{ijgt} = [\, \gamma_1^{low}\; I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 1} \cdot I_{jgt}^{low\ median} \;+\; \gamma_{100}^{low}\; I_{i,jg,t-1}^{decile\ 10} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 10} \cdot I_{jgt}^{low\ median} \,] \;+$$
$$[\, \gamma_1^{medium}\; I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 1} \cdot I_{jgt}^{medium\ median} \;+\; \gamma_{100}^{medium}\; I_{i,jg,t-1}^{decile\ 10} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 10} \cdot I_{jgt}^{medium\ median} \,] \;+$$
$$[\, \gamma_1^{high}\; I_{i,jg,t-1}^{decile\ 1} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 1} \cdot I_{jgt}^{high\ median} \;+\; \gamma_{100}^{high}\; I_{i,jg,t-1}^{decile\ 10} \cdot \bar{I}_{-i,jg,t-1}^{decile\ 10} \cdot I_{jgt}^{high\ median} \,] \;+$$
$$\boldsymbol{I_i^{student}}\,\boldsymbol{\gamma_{101}} \;+\; \boldsymbol{I_g^{grade}} \cdot \boldsymbol{I_t^{school\ year}}\,\boldsymbol{\gamma_{102}} \;+\; \varepsilon_{sgt} \;+\; \varepsilon_{ijgt}$$

We plot coefficient estimates from equation (5) in Figures 3 through 6. Readers should avoid interpreting individual point estimates in these figures. Rather, they should look for patterns that appear relatively consistently. We admit that there is some "art" to the interpretation of these figures, primarily because, by combining the peer effects models with sufficient dexterity, we might explain many patterns. However, we shall be mindful of our results from Table 3 and Figures 1 and 2, in which certain models have already been rejected or at least lacked support.

Examine Figure 3, which shows the effects of peers on students who themselves initially score in the second decile. When such students find themselves in classrooms where the median initial score is high, they clearly benefit most from a boost in the share of classmates who score in the bottom few deciles. They benefit least from a boost in the share of classmates who score in the top few deciles. (The difference in benefit, for a ten percentage point increase in the classmate share, is 6 points or 25 percent of a standard deviation.) When, however, students who are themselves initially low scoring find themselves in classrooms where the median initial score is low, they seem to benefit about equally from classmates of all achievement levels.

Figure 4 shows peer effects for students at the opposite end of the spectrum: those who initially score in the top decile. When such students find themselves in classrooms where the median initial score is high, they benefit most from a boost in the share of classmates who score in the top few deciles. They benefit least from a a boost in the share of classmates who score in bottom deciles. (The difference in benefit, for a ten percentage point increase in the classmate share, is 12 points or 50 percent of a standard deviation.) In contrast, when students who are themselves initially high scoring find themselves in classrooms where the median initial score is low, they benefit most from a boost in the share of classmates who score around the 35th percentile–in other words, close to but a bit above class's median. Finally, when students who are themselves initially high scoring find themselves in classrooms where the median initial score is medium, they benefit as much from a boost in the share of classmates who score around the 45th or 55th percentile as from a boost in the share of classmates who score around the 85th percentile.

What are we to make of Figures 3 and 4? We can explain them with a combination of the Boutique and Focus model along with a general monotonicity property that says that, *all else equal*, a higher achieving peer is better than a lower achieving one. With this combination, Figure 4 makes sense. If a student is initially very high achieving and his classroom has a high median, then he benefits most from peers who are also very high achieving. They, first, reinforce the critical mass at his "Boutique" and, second, drag the median slightly in his direction. This movement of the median means the class's focus shifts slightly in his direction. Yet, there is little chance of bimodality developing, which would cause the "schizophrenia" the Focus model suggests is bad. The same initially high scoring student in a classroom with a low median benefits most from peers who on are his side of the median but not far from it. Their advent moves the class's focus in his direction but they do not generate bimodality. In contrast, the advent of other anomalous, high achieving peers is a mixed blessing. They reinforce his Boutique but, in so doing, generate a distribution that is bimodal and, thus, anti-Focus. Clearly, a combination of the Boutique and Focus models can also explain the intermediate line in Figure 4.

The Boutique and Focus models can also explain the initially low scoring students illustrated by Figure 3, especially if we add the monotonicity property to explain why "low median" and "medium median" lines do not have more of a downward slope. In a low median classroom, the initially low scoring student benefits most from classmates drawn from the first few deciles: they reinforce his Boutique, move the median toward him, and induce no bimodality. In a medium median classroom, he benefits from a boost in the share of students like him, but also in the share of students at the median:

they reinforce classroom focus and are slightly better peers. The high median line is nearly flat with a uptick at each end. The initially low scoring student benefits from other low scoring peers who reinforce his Boutique, from mid scoring peers who pull the median in his direction and are slightly better peers, and from high scoring peers who reinforce class focus and are better peers.

We will let the reader confirm for himself that the combination of Boutique, Focus, and monotonicity can also explain Figures 5 and 6, which are intermediate cases. Figure 5 plots estimated coefficients for students who initially score in the fifth decile; Figure 7 does the same for students who initially score in the eighth decile.

Figures 3 through 6 would be very difficult to reconcile with the Rainbow or Subculture models because each of them implies that augmenting the classroom's focus (on at least some types of students) is a bad. We also see little support for the Bad Apple, Shining Light, or Invidious Comparison models. The monotonicity property is related to the linear-in-means and single-crossing models–suggesting a basic notion on which they are founded makes sense. However, they are clearly rejected as *standalone* models of peer effects.

VI. Do Peer Race, Ethnicity, or Income Matter?

It is far from obvious that, once we have properly accounted for the effects of peers' achievement, peers' race, ethnicity, income, or other characteristics affect a student at all. One hardly knows what to make of statement like the following, quoted in a study of Wake County's desegregation policies:

> [A] high concentration of low-income students . . .appears to have negative effects on students, teachers and the school, and these effects extend beyond the effect of individual students' economic condition.[21]

Are concentrations of poverty bad, in and of themselves, or are they merely proxying for peer

---

[21] Wake County Public School System, Evaluation and Research Department (1999), quoted in Silberman (2003).

achievement for which a researcher has taken insufficient account? Put another way, we have seen that the data consistently rejects the linear-in-means model as a standalone explanation of peer effects. Thus, researchers' common reliance on the linear-in-means model guarantees that any effects of peers that operate non-linearly or through moments other than the mean become omitted variables. These omitted variables will make themselves felt through any available covariate that is correlated with them, and peers' race and income are likely candidates for such covariates.

Thus, we now add indicators of peers' race, ethnicity, income, and other characteristics to the specification (equation (5)) and "run a horse race" to see whether the non-achievement variables matter. For this "horse race," it is useful that Wake County switched its reassignment policy in the middle of the period covered by our data. In the pre-2000 period, students disproportionately experienced reassignment-driven changes in racial composition; in the post-2000 period, students disproportionately experienced reassignment-driven changes in poverty composition. This alteration in the predominant type of "experiments" helps us to identify separately the effects of peers' race, income, and achievement.

The results of interest are shown in Table 4. (We do not show the coefficients on peers' achievement because they are largely unchanged, as will be fairly obvious after we discuss the coefficient estimates for the newly introduced variables.)

The main message of Table 4 is that race, ethnicity, and income do not matter much once we have accounted for the effects of peers' achievement. Twenty-five of the thirty coefficients shown in Table 4 are not statistically significant from zero. Moreover, the coefficients that are statistically significant have magnitudes that are small relative to what would interest a policy maker or relative what to naive studies (like the ones to which the quotation refers) suggest. Consider the few coefficients that are statistically significant. If a student who is himself black and poor experiences a ten percent increase in the share of his class that is black and poor, his achievement falls by 0.6 points (about 2.5 percent of a standard deviation). No other group of students, however, suffers a negative, statistically significant effect when the share of their class that black and poor rises. Indeed, even the point estimates were statistically significant, they are either positive or of such small magnitudes that the effects would be trivial.

If a student who is himself Hispanic and poor experiences a ten percent increase in the share of his class that is Hispanic and poor, his achievement falls by 1.3 points (about 5 percent of a standard deviation). In contrast, if a poor black student experiences a ten percent increase in the share of his class that is Hispanic and poor, his achievement apparently *rises* by 0.8 points–about 4 percent of a standard deviation.

Finally, a student who is himself white or Asian and non-poor sees his achievement *rise* by 0.08 points (0.3 percent of a standard deviation) if the share of his class that is black and non-poor rises by 10 percent. He sees his achievement fall by 0.2 points (0.8 of a standard deviation) if the share of his class that is white or Asian and poor rises by 10 percent.

In short, Table 4 suggests that concentrations of students who black and poor or Hispanic and poor do have negative effects on achievement, but the impacts are small. The vast majority of the apparent impact of a concentration of racial minorities, ethnic minorities, or poor students is really the effect of their achievement. Put another way, if we see two schools with the same *distribution* of achievement (not merely the same mean), we should expect their students' achievement to evolve similarly in the future, even if the schools have quite different racial, ethnic, and income compositions. Of course, policy makers might still wish to equalize the two schools' racial, ethnic, and income compositions for purely social reasons.

Table 4 suggests that Wake County's policy switch, evaluated as written, was (very slightly) good for students who were poor and black or Hispanic and was (very slightly) bad for students who were non-poor and white or Asian. The former result is because, with the policy change, poor black and poor Hispanic students should have gained non-poor black and non-poor Hispanic peers and lost poor peers of all races. Since concentrations of poor blacks and poor Hispanics have a negative effect on achievement, the overall impact is positive. The latter result is because, with the policy change, non-poor white and Asian students should have lost non-poor black peers (who are good for their achievement) and gained poor white and Asian peers (who are bad for their achievement). Nevertheless, the overall conclusion should be that switching policies, *per se*, had little effect.

Our results suggest that greater effects were probably intentionally induced by reassignments that shook up the distribution of peer achievement in schools. For instance, our results suggest that reassignments were beneficial if they created schools in which there was a critical mass of students at each achievement level represented in the school. On the other hand, reassignments were pernicious if they created schools whose children have bimodal or simply very diffuse achievement distributions.

In addition to looking for independent effects of race, ethnicity, and income, we investigated whether the sex composition or parental education composition of peers mattered. We found ample evidence that girls are more beneficial peers than boys, even after taking account of achievement, race, ethnicity, and income. This finding confirms the results of previous work by the author (Hoxby 2000), a study with an empirical strategy more attuned to analyzing the effects of peer sex composition. Rather unexpectedly, we did not obtain any evidence that students whose parents are more educated make more beneficial peers. We surmise that parents' education does not have an independent effect once we have taken account of peers' achievement.

VI. Conclusions

Our most important findings are three. First, certain very commonly employed models of peer effects, such as the linear-in-means and single-crossing models, are rejected as standalone models. In other words, they do not sufficiently embody peer effects to be used, by themselves, to generate empirical specifications. Of course, because we find general support for the notion that higher achieving people are better peers all else equal, the linear-in-means and single-crossing models may still inspire *parts* of an adequate empirical specification. Such a specification should also be able to embody other models of peer effects, especially the Boutique and Focus models.

Second, our finding support for the Boutique and Focus models suggests that schools, colleges, and workplaces should be wary of creating peer groups in which some people are isolated (in terms of prior achievement, innate ability, or productivity). However, they should also avoid creating critical mass around a certain type of person if, by so doing, they generate a peer group that is bimodal or, more generally, multimodal. Some focus is good. Our finding support for the Boutique and Focus models also suggests that real-world stratification across schools, colleges, neighborhoods, workplaces, and metropolitan areas is probably not generated by the Single-Crossing Model (the main appeal of which has always been its mathematical elegance, anyway). As a result, we may want to revisit models of school choice, college choice, and urban economics that rely heavily on the single-crossing assumption.

Notice that our evidence does not suggest that complete segregation of people, by types, is optimal. This is because (a) people do appear to benefit from interacting with peers of a higher type and (b) people who are themselves high types appear to receive sufficient benefit from interacting with peers a bit below them that there is little reason to isolate them completely. What our evidence *does* suggest is that efforts to create interactions between lower and higher types ought to maintain continuity of types.

Finally, we find strong evidence that peers' race, ethnicity, and income have only very slight effects once we have properly accounted for peers' achievement. This suggests that fears of racial, ethnic, and economic desegregation are overblown; but it also suggests that policy makers who pin all their hopes for achievement on such desegregation are unduly optimistic. In conducting racial, ethnic, and economic desegregation, policy makers ought to pay more attention to how they are affecting the distribution of achievement within peer groups. The distribution of achievement should probably be of primary concern, not an unintended consequence.

References

Angrist, Joshua D. and Lang, Kevin. (2002). "How Important are Classroom Peer Effects? Evidence from Boston's Metco Program." NBER Working Paper 9263.

Betts, Julian R. and Zau A. (2004). "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data." Unpublished manuscript.

Benabou, Roland. (1996) "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, 86:3, pp. 584-609.

Boozer, Michael A. and Cacciola, S.E. (2001). "Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data." Unpublished manuscript.

Brock, William A. and Durlauf, Steven N. (2001) "Interactions-Based Models," Chapter 54 in eds. James Heckman and Edward Leamer, *Handbook of Econometrics*, Volume 5. Amsterdam: Elsevier Science B.V., pp. 3297-3380.

Carrell, Scott E., Malmstrom, Frederick V., and West, James E. (2005) "Peer Effects in Academic Cheating," Unpublished manuscript.

Cascio, Elizabeth, Gordon, Nora, Lewis, Ethan, and Reber, Sarah J. (2005) "Financial Incentives and the Desegregation of Southern Public Schools," Unpublished manuscript.

Clotfelter, Charles T. (2004) *After Brown: The Rise and Retreat of School Desegregation*. Princeton, NJ: Princeton University Press.

Epple, Dennis, Richard Romano, and Holger Sieg. (2003). "Peer Effects, Financial Aid, and Selection of Students into Colleges," *Journal of Applied Econometrics*.

Glaeser, Edward L., Sacerdote, Bruce L., and Scheinkman, Jose A. (2003). "The Social Multiplier," *Journal of the European Economic Association*, 1, 345 - 353.

Graham, Bryan S. (2004). "Identifying Social Interactions through Excess Variance Contrasts." Unpublished manuscript.

Hoxby, Caroline M. (2000). "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Paper 7867.

Kremer, Michael. (1993) "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, August, pp. 551-576.

Kremer, Michael, and Levy, Daniel. (2003). "Peer Effects and Alcohol Use among College Students," Unpublished manuscript.

Lazear, Edward. (2001) "Education Production," *Quarterly Journal of Economics*, Vol. 116.3 (August), pp. 777-803.

Manski, Charles F. (1993). "Identification and Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60, 531-542.

Nechyba, Thomas. (1996) "Public School Finance in a General Equilibrium Tiebout World: Equilization Programs, Peer Effects, and Vouchers." NBER Working Paper 5642.

Ogletree, Charles. (2004) *All Deliberate Speed: Reflections on the First Half Century of Brown V. Board of Education*. New York: W.W. Norton & Co.

Reber, Sarah J. (forthcoming) “Court-Ordered Desegregation: Successes and Failures in Integration Since Brown,” *Journal of Human Resources*.

Sacerdote, Bruce L. (2001). "Peer Effects with Random Assignment: Results for Dartmouth Roomates." *Quarterly Journal of Economics*, 116, 681-704.

Samms, Gavin. (2004) "Desegregation, Peer Effects, and Achievement: Evidence from a Policy Experiment." Unpublished manuscript.

Silberman, Todd. (2003). *Wake County Schools: A Question of Balance*. Unpublished monograph sponsored by The Spencer Foundation, Raleigh, N.C.

Stinebrickner, Todd R. and Stinebrickner, Ralph. (2001). "Peer Effects Among Students from Disadvantaged Backgrounds," University of Western Ontario, CIBC Human Capital and Productivity Project Working Paper No 20013.

Vidgor, Jacob and Nechyba, Thomas. (2004). "Peer Effects in North Carolina Public Schools." Unpublished manuscript.

Wake County Public School System, Evaluation and Research Department. (1999) "The Impact of Poverty Upon Schools," Evaluation and Report Report No. 99.20, Marchs, Raleigh, N.C.

Weingarth, Gretchen. (2005) *Taking Race Out of the Equation: The Effect of Changing Classroom Poverty Concentrations on Student Achievement*. Harvard University Senior Honors Thesis in Economics. Cambridge: Harvard University Archives.

Zimmerman, David J. (2003). "Peer Effects in Academic Outcomes: Evidence From a Natural Experiment." *The Review of Economics and Statistics*, 85,1, 9–23.

Zimmerman, David J. and Winston, Gordon. (2004). "Peer Effects in Higher Education," in ed. Caroline Hoxby, *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*. Chicago: University of Chicago Press.

Figure 1

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

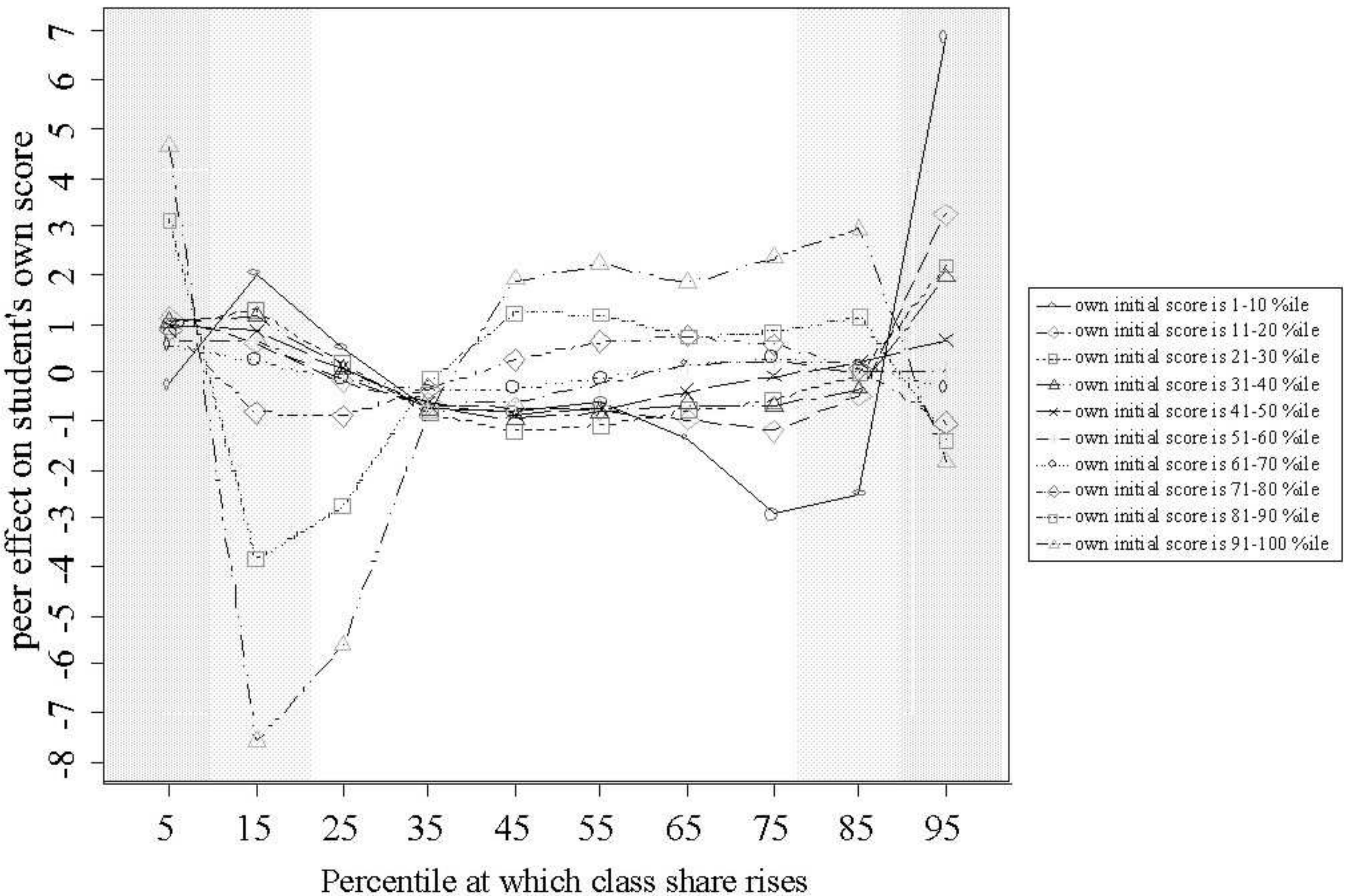

Figure 2

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

effects for various initial scores

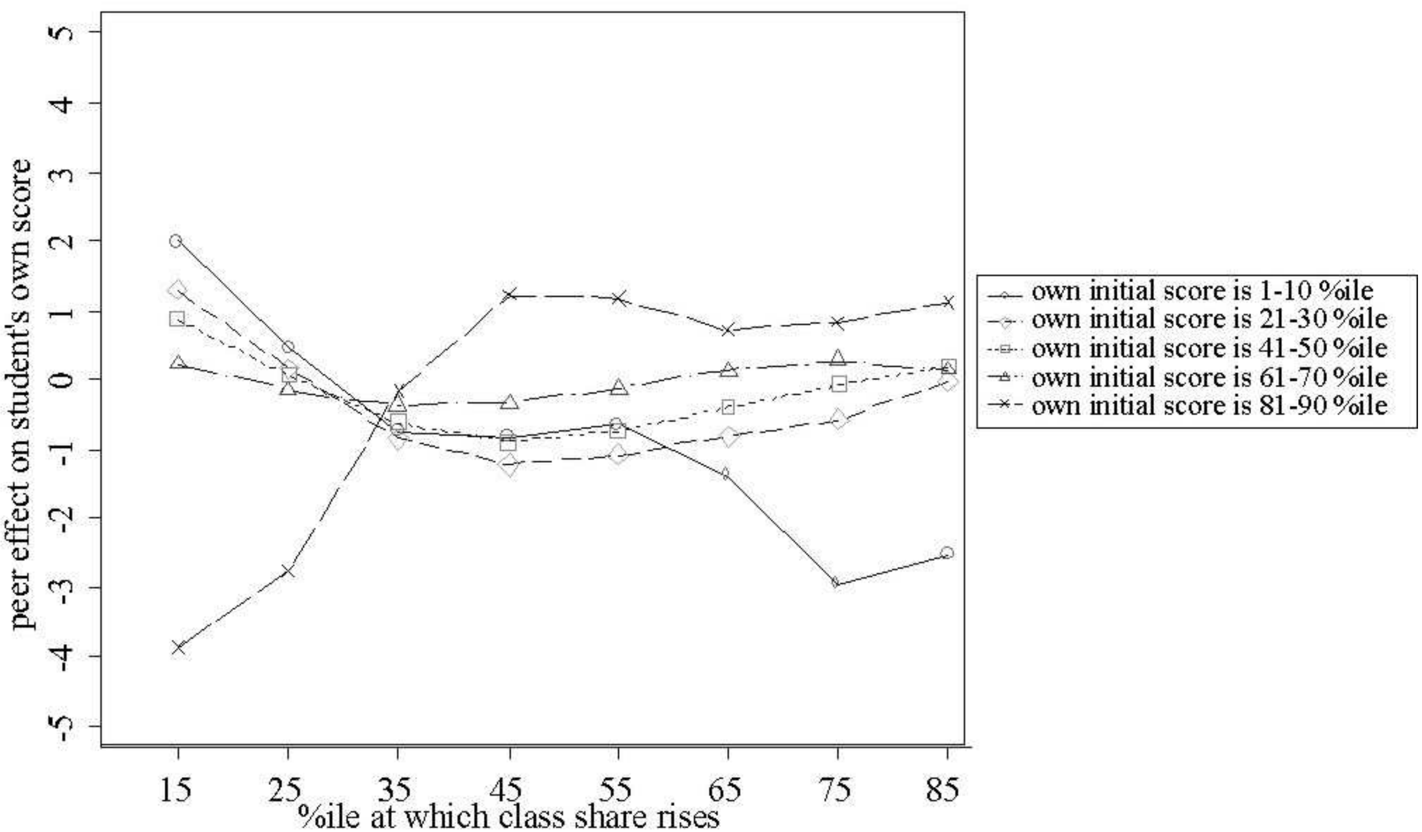

Figure 3

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

students whose own initial score is 11-20 %ile

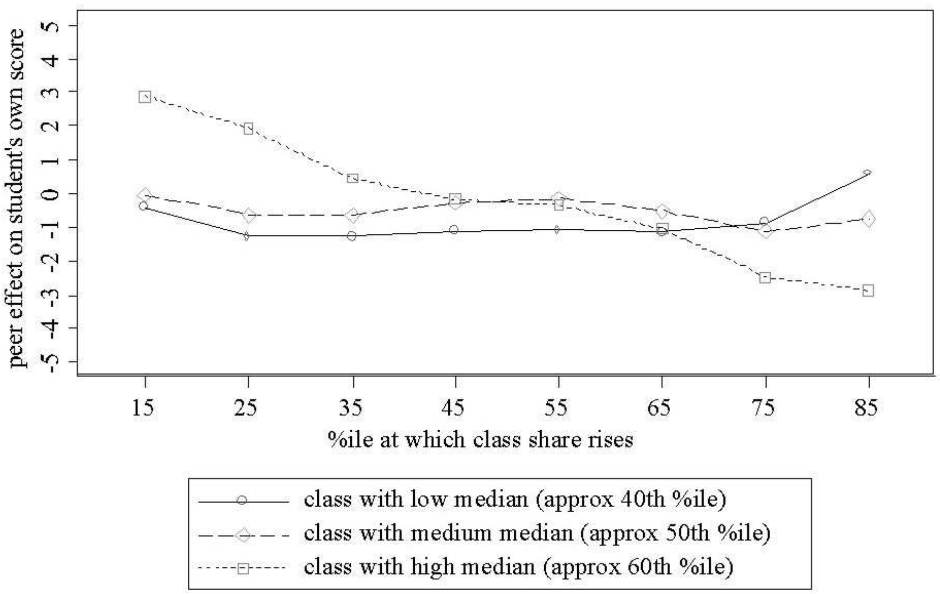

Figure 4

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

students whose own initial score is 91-100 %ile

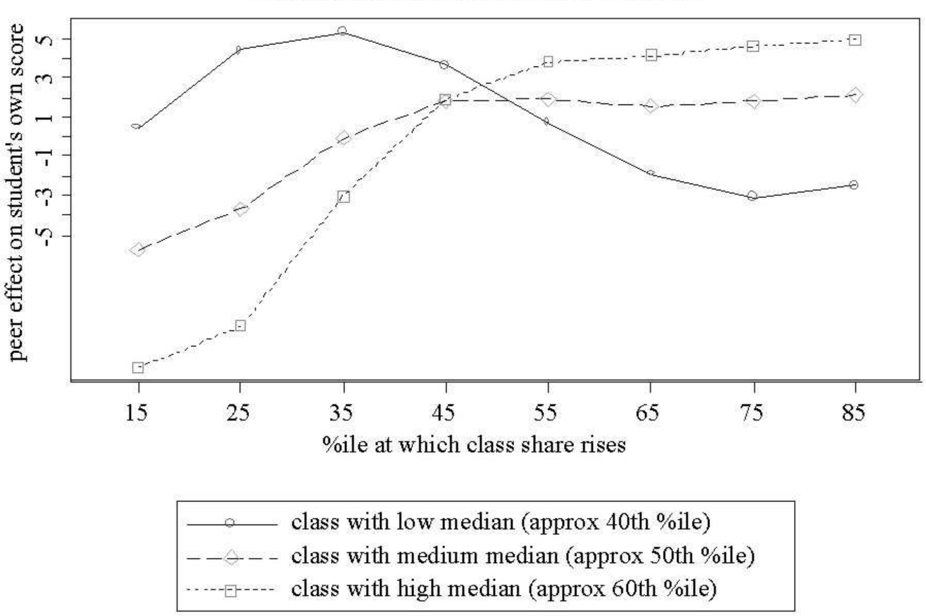

Figure 5

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

students whose own initial score is 41-50 %ile

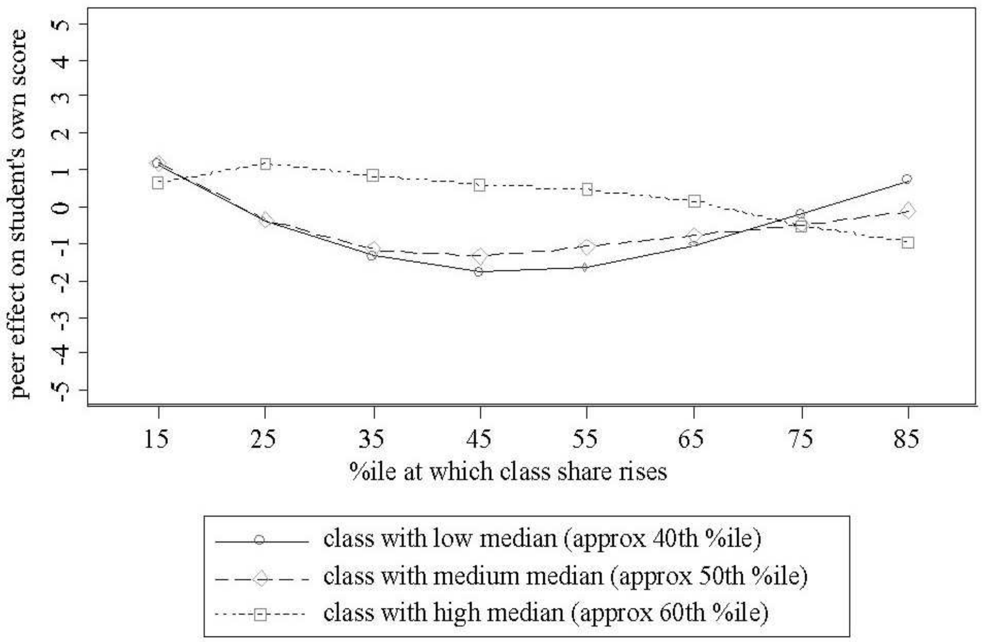

Figure 6

Effect of Raising Share of Class at Various Achievement Levels by 10 Percent

students whose own initial score is 71-80 %ile

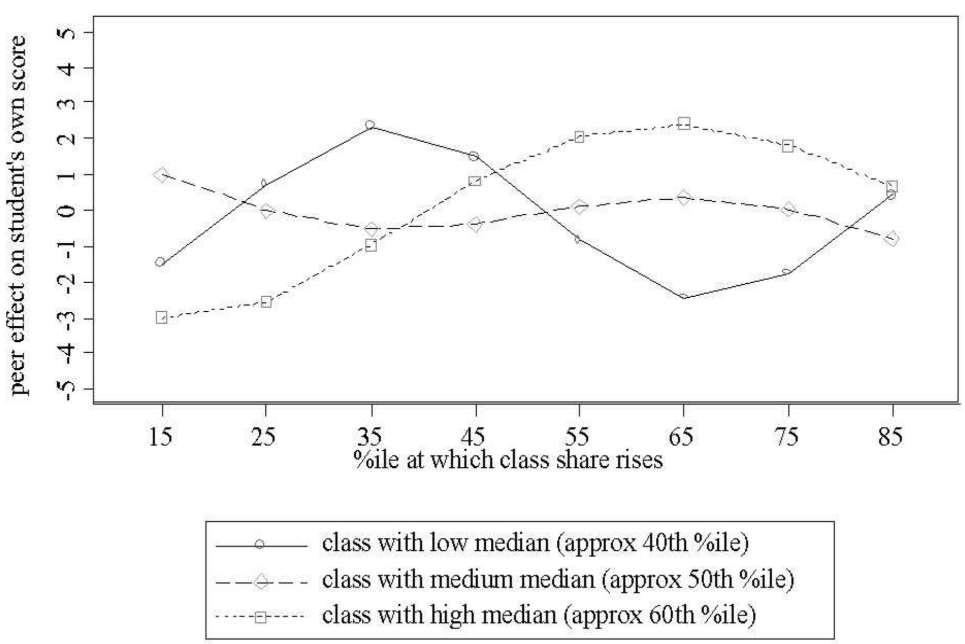

Table 1
Descriptive Statistics for the Wake County Dataset

Time-Constant Variables (one observation per student)[a]			
Female	131138	0.489	0.500
Black	131138	0.262	0.440
White or Asian	131138	0.671	0.470
Hispanic, Mixed, or "Other Race"	131138	0.067	0.249
Parents' Education: Less than High School	131138	0.057	0.232
Parents' Education: High School Diploma or Equivalent	131138	0.255	0.436
Parents' Education: Some Postsecondary but No Degree	131138	0.140	0.347
Parents' Education: Two Year College	131138	0.364	0.481
Parents' Education: Four Year College	131138	0.169	0.375
Parents' Education: Graduate School	131138	0.015	0.122
Initial Total Scale Score (de-meaned by grade·year)[b]	131138	-0.737	24.424
Ever Experienced a Policy-Driven Change in Peer Composition	131138	0.624	0.484
Was Ever Reassigned by Policy	131138	0.239	0.426
Time-Varying Variables (multiple observations per student)[c]			
Reading Scale Score (pre-2003 scale)	357358	157.2	10.3
Math Scale Score (pre-2001 scale)	357358	163.0	15.5
Total Scale Score (de-meaned by grade·year)	357358	0.000	23.081
Grade	357358	5.365	1.713
Spring of School Year	357358	1999	2
Size of Cohort	357358	228	128
Learning Disabled	357358	0.073	0.261
Other Disability (Individual Education Program)	357358	0.070	0.256
Participate in Free Lunch (1998-99 onwards)	237867	0.171	0.377
Participate in Reduced-Price Lunch (1998-99 onwards)	237867	0.048	0.214
Class's Previous Period Mean Reading Score	245780	154	9
Class's Previous Mean Math Score	245780	157	14
Share of Class with Previous Score between 1 and 10th Percentiles	245780	0.101	0.141
Share of Class with Previous Score between 10 and 20th Percentiles	245780	0.102	0.100
Share of Class with Previous Score between 20 and 30th Percentiles	245780	0.097	0.084
Share of Class with Previous Score between 30 and 40th Percentiles	245780	0.095	0.076
Share of Class with Previous Score between 40 and 50th Percentiles	245780	0.096	0.073
Share of Class with Previous Score between 50 and 60th Percentiles	245780	0.095	0.072
Share of Class with Previous Score between 60 and 70th Percentiles	245780	0.098	0.074
Share of Class with Previous Score between 70 and 80th Percentiles	245780	0.098	0.078
Share of Class with Previous Score between 80 and 90th Percentiles	245780	0.100	0.087
Share of Class with Previous Score between 90 and 100th Percentiles	245780	0.118	0.125
Simulated Instrument Cohort's Initial Mean Reading Score	245780	155	8
Simulated Instrument Cohort's Initial Mean Math Score	245780	159	12
Share of Simulated Cohort with Initial Score between 1 and 10th Percentiles	245780	0.120	0.075
Share of Simulated Cohort with Initial Score between 10 and 20th Percentiles	245780	0.104	0.051
Share of Simulated Cohort with Initial Score between 20 and 30th Percentiles	245780	0.097	0.039
Share of Simulated Cohort with Initial Score between 30 and 40th Percentiles	245780	0.093	0.038
Share of Simulated Cohort with Initial Score between 40 and 50th Percentiles	245780	0.093	0.034
Share of Simulated Cohort with Initial Score between 50 and 60th Percentiles	245780	0.095	0.036
Share of Simulated Cohort with Initial Score between 60 and 70th Percentiles	245780	0.095	0.034
Share of Simulated Cohort with Initial Score between 70 and 80th Percentiles	245780	0.096	0.038
Share of Simulated Cohort with Initial Score between 80 and 90th Percentiles	245780	0.097	0.048
Share of Simulated Cohort with Initial Score between 90 and 100th Percentiles	245780	0.115	0.073

table notes may be found on the next page

Table 1

Descriptive Statistics for the Wake County Dataset

Notes:

[a] A student is included in the dataset if we ever observe his or her end-of-grade test scores.

[b] The demeaned total scale scores are the residuals from a linear regressions of students' scale scores on an exhaustive set of grade-by-school year indicators.

[c] We observe 64,785 in one year only; 45,950 students in two years; 57,702 in three; 43,336 in four; 46,540 in five, 106,248 in six; 7,364 in seven; and 504 in eight years. These numbers include students who have missing test scores in one or more years. A student who is making his first appearance in the dataset has a missing observation for the class and simulated cohort variables. In the analyses, we impute free and reduced-lunch status for the school years before 1998-99 by backcasting a student's later status and filling in the remaining missing observations using a prediction based on parents' education.

Source: Authors' calculations based on Wake County data from the North Carolina Education Research Data Center.

Table 2

Tests of Whether Experiencing Policy-Driven Changes in Peers is a Function of Student's Own Characteristics (apart from race, free or reduced-price lunch, and other factors considered in reassignment)

	Dependent Variable	
	Experienced a Policy-Driven Change in Own Cohort	Reassigned
Initial Test Score (total scale score de-meaned by grade·year)	1.33E-05	-3.10E-05
	(1.11E-05)	(3.23E-05)
Parents' Education: Less than High School	0.0032	-0.0317
	(0.0137)	(0.0484)
Parents' Education: High School Diploma or Equivalent	0.0028	-0.0283
	(0.0137)	(0.0484)
Parents' Education: Some Postsecondary but No Degree	0.0026	-0.0320
	(0.0137)	(0.0484)
Parents' Education: Two Year College	0.0024	-0.0232
	(0.0137)	(0.0484)
Parents' Education: Four Year College	0.0028	-0.0193
	(0.0138)	(0.0487)
Parents' Education: Graduate School	0.0023	-0.0293
	(0.0137)	(0.0484)
Race and Ethnicity Indicators	yes	yes
Free and Reduced-Price Lunch Indicators	yes	yes
Grade-by-School Year Effects	yes	yes
Initial School Effects	yes	yes

Notes: The table shows estimated coefficients from two linear probability models with the dependent variables listed. Standard errors are in parentheses. The idea is to test whether being "treated" with policy-driven changes in peers is a function of variables other than those explicitly considered by the reassignment authorities. If we were to find evidence that the authorities were discriminating among students (with regard to reassignment) along dimensions they were not supposed to consider, it would suggest that treatment was not random conditional on a student's fixed characteristics. For descriptive statistics on the variables and data source, see Table 1.

Table 3
Effects of Peers on Student's Own Score,
Linear-in-Means Specification and Other Homogeneous Treatment Effect Specifications
Dependent Variable: Student's Own Test Score[a]

	Least squares	Simulated instrumental variables (simulated cohort)[c]		
Class's initial mean test score[b]	-0.002 (0.002)	**0.254** **(0.092)**	**0.351** **(0.115)**	**0.348** **(0.135)**
Share of class with initial test score below 25th percentile[b]			-1.958 (4.771)	
Share of class with initial test score 25th & 50th percentiles[b]			**-32.067** **(5.700)**	
Share of class with initial test score above 75th percentile[b]			**-13.499** **(3.649)**	
Share of class with initial test score 10th & 20th percentiles[b]				**21.248** **(11.537)**
Share of class with initial test score 20th & 30th percentiles[b]				0.912 (6.143)
Share of class with initial test score 30th & 40th percentiles[b]				**-18.260** **(13.943)**
Share of class with initial test score 40th & 50th percentiles[b]				-7.595 (10.983)
Share of class with initial test score 50th & 60th percentiles[b]				11.043 (11.545)
Share of class with initial test score 60th & 70th percentiles[b]				**18.179** **(11.947)**
Share of class with initial test score 70th & 80th percentiles[b]				**24.666** **(10.109)**
Share of class with initial test score 80th & 90th percentiles[b]				3.692 (9.668)
Share of class with initial test score above 90th percentile[b]				-7.907 (10.001)
Grade-by-school year effects	yes	yes	yes	yes
School Effects	yes	yes	yes	yes
Student Effects	yes	yes	yes	yes

Table notes continue on next page.

Table 3
Effects of Peers on Student's Own Score,
Linear-in-Means Specification and Other Homogeneous Treatment Effect Specifications
Dependent Variable: Student's Own Test Score[a]

Notes:
[a] Student's test score is the sum of his math scale score and reading scale score.
[b] Class is always class excluding student himself. Similarly, cohort is always cohort excluding student himself. A cohort is a school by grade by school year group of students–for instance, third graders in school X in the 1999-00 school year.
[c] The simulated cohort is the cohort the student would have experienced if reassignments (only) had taken place but all potentially endogenous peer moves were disallowed. See text for additional detail.

Additional Notes: The table shows estimated coefficients from linear regressions and instrumental variables regressions with the dependent variables listed. Standard errors are in parentheses. For descriptive statistics on the variables and data source, see Table 1.

Table 4
Effects of Peers' Race, Income, and Other Characteristics on Student's Test Score,
(peers' achievement is included via the heterogeneous treatment effect specification: equation (4) by class's median score)

Dependent Variable: Test Score[a] of a Student who is...

	Black and Poor	Black and Non-Poor	Hispanic[c] and Poor	Hispanic[c] and Non-Poor	White and Poor	White and Non-Poor
Share of class that is black and poor[b]	**-6.127**	-0.364	1.979	-0.469	2.769	0.634
	(1.464)	(0.856)	(2.944)	(2.630)	(2.126)	(0.477)
Share of class that is black and non-poor[b]	-0.778	-0.382	0.879	-1.381	-1.439	**0.772**
	(1.011)	(0.487)	(2.289)	(1.751)	(1.853)	**(0.337)**
Share of class that is Hispanic[c] and poor[b]	**8.129**	3.104	**-13.311**	-6.598	-3.962	0.020
	(3.757)	(2.889)	**(8.410)**	(7.012)	(5.977)	(1.431)
Share of class that is Hispanic[c] and non-poor[b]	1.487	1.836	11.156	4.741	-4.377	0.838
	(6.207)	(3.744)	(11.218)	(9.817)	(9.222)	(1.648)
Share of class that is white or Asian and poor[b]	-0.235	-0.554	5.258	2.701	-1.506	**-2.117**
	(2.843)	(2.440)	(6.837)	(6.272)	(4.196)	**(1.234)**
Peers' achievement, heterogeneous treatment effect specification: equation (4) by class's median score	yes	yes	yes	yes	yes	yes
Grade-by-school year effects	yes	yes	yes	yes	yes	yes
School Effects	yes	yes	yes	yes	yes	yes
Student Effects	yes	yes	yes	yes	yes	yes

[a] Student's test score is the sum of his math scale score and reading scale score.
[b] Class is always class excluding student himself. Similarly, cohort is always cohort excluding student himself. A cohort is a school by grade by school year group of students–for instance, third graders in school X in the 1999-00 school year.
[c] "Hispanic" is actually Hispanic ethnicity, mixed race, or other race.

Notes: The table shows estimated coefficients from simulated instrumental variables regressions with the dependent variables listed. Standard errors are in parentheses. The simulated instruments are based on the cohort the student would have experienced if reassignments (only) had taken place but all potentially endogenous peer moves were disallowed. See text for additional detail. For descriptive statistics on the variables and data source, see Table 1.