Parent Choice Versus Attendance Area Assignment to Schools:
Does Magnet-Based School Choice Affect NAEP Scores?

Douglas A. Archbald & David Kaplan
University of Delaware

Abstract: Inter- and intra-district public school choice, vouchers, tuition tax credits and other forms of school choice have been advocated for decades, in large part on grounds that the market forces engendered will improve public education. There are many studies of school choice policies and programs and a large theoretical literature on school choice, but thus far no studies have used a large national sample and common metric to perform a multi-level, multi-district analysis of relationships between school choice policy and student achievement. This study links a national sample of NAEP student achievement data, with district level information on magnet-based school choice policy, and with demographic data from the U.S. Census. Using three-level hierarchical linear modeling we find substantial effects of school and district demographic variables on student achievement, but, after adjusting for multi-level demographic characteristics, we find only small differences in student achievement in school districts with magnet schools and school choice policies as compared with districts with attendance area based student assignment, and no magnet schools. NAEP achievement scores

Address all correspondence to: Douglas A. Archbald, School of Education, University of Delaware, 113 Willard Hall, Newark, Delaware 19716-2922. E-mail: archbald@udel.edu

International Journal of Educational Policy, Research, & Practice, Volume 5, Number 1, Spring 2004. ISSN 1528-3534;

are marginally lower in the sample of students within districts reporting magnet schools and associated school choice policies.

Introduction

School choice in a variety of forms has expanded nationally and internationally, in large part on the belief that market forces can improve public education (Peterson 2001; Teelken, 1999; Van Zanten, 1996; Whitty, Power, & Halpin, 1998). In the United States, magnet schools were the first school choice policy widely implemented in public education (Blank, Dentler, Baltzell & Chabotar, 1983; Levine & Havighurst, 1977). Researchers have studied magnet schools to learn whether this approach to school choice affects achievement, but in almost all cases studies have focused on just a single or a few districts. This study merges existing national data sets to perform a multi-level, multi-district analysis. It charts new ground by linking student achievement data from the National Assessment of Education Progress (NAEP) with district level policy information on school choice and demographic data from the U.S. Census.

Of the many different forms of school choice in existence (or proposed), magnet schools implement a regulated form of school choice—neither private schools nor charter schools are involved, and parents' choices may be restricted in the interests of racial desegregation of schools. Nonetheless, the presence of magnet schools in a district still produces a great deal of parental choice, raising questions about whether school districts with magnet-based school choice might have higher achievement than districts lacking school choice. This is the question we examine.

Theory and Research on Magnet Schools, School Choice, and Student Achievement

The Growth of Magnet Schools and School Choice

While charter schools and voucher programs have been getting more attention recently, magnet schools remain one of the most common approaches in the U.S. (CER, 1998; Steel & Levine, 1994). The "magnet school boom" (Warren, 1978) of the late 70s and continuing into the 80s was the precursor of the widespread adoption of public school choice policies we see today. Since 1976 when only 14 districts had magnet schools (Blank, Dentler, Baltzell & Chabotar, 1983), the number of districts with magnet schools has grown rapidly along with enrollment in magnet schools (Blank, 1989). By 1990 estimates were that almost half of urban districts had magnet schools (Steel & Levine, 1994).

When a district implements a magnet-based school choice policy, it designates a portion of the schools in the district as magnet schools. Magnet schools are open to enrollment by choice on a district-wide basis, subject to racial balance guidelines (sometimes regions within districts are used if the district is geographically large). Magnet schools typically have special features, programs, or instructional methods that distinguish them from "traditional" schools. Magnet schools may offer distinctive instructional approaches such as Montessori or Waldorf; curricular emphases such as art, math, science, technology, or foreign language immersion; or disciplinary climates involving above average regimentation and required uniforms.[1] In addition, magnet schools sometimes offer special features like smaller class sizes, instructional aides, or newer facilities, features made possible by federal funds from the Magnet Schools Assistance Program that has supported magnet schools' racial integration function since the mid-70s. Districts with magnet schools often extend choices to cover other schools in the district as well. Parents may choose and enroll a child in any school as long as the choice does not increase racial desegregation in the district. A child in a school where his/her race is in the majority can transfer to any school where his/her race is in the minority.

Theory and Research on Magnet-based School Choice

A large academic and practitioner-oriented literature supported the growth of magnet schools and public school choice. Magnet-based school choice was supported because it helped voluntary racial integration (e.g., Greeley, 1987; Levine & Havighurst, 1977; Rossell, 1990), expanded educational alternatives for parents, and stimulated competition in public schools (Barr, 1982; Blank, 1989; Clinchy & Cody, 1978; Finn, 1985; Finn, 1987; Kolderie, 1985; Nathan, 1989; Manley-Casimir, 1982; Raywid, 1985; Toch, Linnon, & Cooper, 1991, May 27).

Researchers and policymakers have long been interested in achievement effects associated with magnet schools (e.g., Archbald, 1995; Blank, 1989; Blank & Archbald, 1992; Gamoran, 1996; Goldhaber, 1999; Larson, Witte, Staib, & Powell, 1990). Two inter-related arguments have supported the growth of magnet schools, both suggesting achievement benefits. One argument is that the expansion of educational alternatives improves the fit between a school district's educational program and its clientele's preferences. Michaelson (1981) has referred to this as "allocative efficiency." It reflects the assumption that not all parents want the same thing, and so rather than force all parents to accept the same kind of school program, school systems should have differentiation among schools reflecting different educational philosophies, pedagogies, and curriculum specializations. In theory, magnet schools do this. While they may not be for everyone, they make it possible for a school district to accommodate more parents with varied preferences than would otherwise be possible with a "one size fits all" philosophy. This viewpoint is prominent in the literature supporting magnet schools (e.g., Fliegel, 1993; Clewell & Joy, 1993; Nathan, 1989; Raywid, 1985).

The second argument reflects free market principles. Advocates of magnet schools have argued that they foster healthy competition in school systems. Schools that have long been assured a designated attendance area clientele, are in a more competitive position when a school system implements magnet-based school choice policies. Children from the neighborhood are no longer guaranteed to a school. When the market is opened up, as the theory goes, schools will try harder to satisfy parents. Observers disagree whether "trying harder" necessarily results in improving curriculum, instruction, and school discipline in ways that raise academic achievement. Still, these two inter-related arguments have been very influential in expanding the adoption of public school choice, and still propel the growth of charter schools and voucher programs.

Ideally, one would want to compare achievement in a sample of districts with magnet-based school choice over time to a matched sample of districts without these policies, but this has not been done for a variety of reasons, not least of which are the methodological difficulties such a study would pose. One research method compares magnet students' achievement to the achievement of other students in the district not in magnet schools. These comparisons typically find magnet school students have higher test scores.For instance, Archbald (1995) compared achievement scores of elementary magnet students, neighborhood school students, and students who had "choiced out" of a neighborhood school, but not to a magnet. The results showed a statistically significant effect of magnet school enrollment, although the effect was fairly small (about 3 NCE points on average). Larson et al. (1990) compared high school magnet students to a sample of nonmagnet students on mathematics and science tests. Since it was known from unadjusted cross-sectional comparisons that the magnet students had higher test scores (on average) and that several of the magnet programs were selective in their admissions,[3] the comparison group was randomly sampled from college-prep courses in the nonmagnet high schools. Their analyses found modestly, statistically significantly, higher mean test scores in the magnet programs, even when adjusting for prior eighth grade achievement.

An important methodological limitation of the above magnet-nonmagnet comparisons is selection bias. Ideally, students should be randomly assigned to magnet and nonmagnet schools, but, of course, students are in their schools based on their own and their parents' decisions, and so the two groups are possibly different on unmeasured covariates of the characteristic that distinguishes the choosers from the nonchoosers. If these covariates are associated with student achievement, then this confounds the magnet-nonmagnet comparisons. Studies of differences between choosing and nonchoosing parents in school choice systems find that parents who take advantage of school choice opportunities often differ in ways that would predict higher achievement among their children (Archbald, 1996; Levine, 1975; Smrekar & Goldring, 1999; Wells, 1991; Witte, 1998).[4]

Gamoran (1996) conducted a large national-sample study comparing magnet and nonmagnet schools using the National Educational Longitudinal Survey (NCES, 1988). Gamoran focused just on urban districts and compared schools that identified themselves as a magnet schools with schools that were not identified with any specialty designation. Among other things, Gamoran found students who attended magnet schools had higher reading and social studies achievement scores, after controlling for pre-existing differences among students.[5]

The above studies, done in the 90s represent the best research on magnet school achievement effects. There have been a few other, older studies of magnet school systems as reviewed in Blank and Archbald (1992), but these studies had such serious methodological shortcomings, usually no comparison groups or demographic adjustments, that the authors concluded "little is known about the effects of magnet schools on the quality of education in our schools" (p. 85).

Research on Other Forms of School Choice

The magnet-based form of school choice examined in this study is among the more regulated approaches. Many researchers and school choice advocates believe the monopolistic structure of public education must be forced to change through stronger free market mechanisms, and are skeptical that the regulated magnet-based choice policies implement market forces sufficient to cause under-performing systems to improve (Chubb & Moe, 1990; Erickson, 1982; Everhart, 1982; Friedman & Friedman, 1981; Harmer, 1994; Hoxby, 1996; Lieberman, 1986; Michaelson, 1981; Moe, 2001; Peterson, 1990; Peterson 2001). Two key differences separate magnet-based school choice from the less regulated charter school and voucher policies. One is the racial balance goals regulating admissions to magnet schools and other participating choice schools in systems with magnets (choice is allowed, but restricted to inhibit resegregation); the other difference is that the magnet-based approach to choice is still "within the system"—that is, the district bureaucracy and teacher association regulations still exercise substantial control over the management and operation of all schools.

Charter schools and voucher eligible private schools are considerably more independent and thus, in theory, create a more open marketplace of school options. Charter schools receive state funding in proportion to their enrollment to fund operating costs, but are essentially independently managed organizations. Many charter schools are even operated through private enterprise arrangements (GAO, 2002). In voucher programs, it is the private schools that participate in the market and because tuition vouchers greatly expand access to private schools, they—again, in theory—create greater inter and intra sector competition.

It is beyond the scope of this paper to review all the research on these different forms of school choice other than note that the existing research has not converged toward definitive conclusions (GAO, 2002; Gill, 2000; Goldhaber, 1999). One body of research stems from evaluations of a handful of pioneering voucher programs in the U.S.—in Milwaukee, Dayton (Ohio), New York City, and Washington, D.C. (Green, Peterson & Du, 1998; Howell, Wolf, Peterson, & Campbell, 2000; Rouse, 1998; Witte, 1998).[6] These evaluations have been methodologically sophisticated, even incorporating random assignment principles. Green et al. (1998) and Howell et al. (2000) discovered significant achievement benefits accruing to the children who received vouchers and attended private schools; however, Rouse (1998) and Witte (1998), found much smaller and in some analyses no positive effects. One of the main methodological shortcomings of the studies finding significant positive effects is differential attrition of subjects from the different conditions over time. A key concern was the loss of apparently less academically competent students from the private school condition—that is, it appeared that there were a number of students awarded vouchers who attended the private schools but later returned to the public schools, and these on average were less academically successful students.[7] Commenting on the state of evidence on educational vouchers, Levin (1998:3) writes, "…both advocates and detractors tend to argue about the consequences of vouchers more from theoretical and ideological grounds than empirical ones." And on the Milwaukee voucher studies, Levin adds, "Even Rouse [Rouse (1998)], the most sophisticated [of the] analyses of these data, is riddled with cautions about data gaps and the assumptions that are made to address them including the use of instrumental variables." Methodological obstacles to definitive research on this question are many, the available sites for research remain limited, and ideological interests are strong.

Research on charter school effects is just now emerging, but there are only a few achievement-effects studies using controlled charter versus non-charter school comparisons (Gill et al., 2001; Miron & Nelson, 2001; 1999; Lin, 2002). A recent review noted the paucity of well designed research (despite a voluminous literature on charter schools) and concluded that, "…evidence on the academic effectiveness of charter schools is mixed" (Gill et al., 2001, p.95). According to Miron and Nelson's review (2001, p. 1) "Overall, the charter impact on student achievement appears to be mixed or very slightly positive. However, this conclusion is tempered by the fact that there are, as yet, no systematic studies of charter achievement in several states that have large numbers of charter schools."

Summary Comments on the Theory and Research

There are many who believe the forces of markets and competition can improve public education. This presumption has supported the growth and development of magnet schools, inter-district public school choice, charter schools, private management of public schools, and private school voucher programs and policies. Some supporters of school choice advocate public-only school choice models. They believe approaches like magnet schools or inter-district school choice policies can create adequate educational alternatives and competition among public schools. However, as described above, a large literature and advocacy group developed in the late 80s and 90s skeptical of the prospects of school choice confined within and managed by the public education bureaucracy. Illuminating studies have emerged on relationships between school choice and academic achievement, but much more research needs to be done because of the many varieties of school choice, the multiple levels at which effects need to be investigated, and the many challenges in measuring achievement outcomes.

This study provides a unique contribution in several ways. It uses a large national sample of districts, but overcomes the problem of different districts using different tests by relying on one common assessment: NAEP. Another important features of this study is that it is not a within-district study comparing magnet schools to nonmagnets from the same district, as so often has been done; rather it asks and investigates the question: do *school systems* with magnet-based school choice have higher NAEP achievement than those without school choice, that is, with attendance zones and centrally administered student assignment to schools.

Method

A Multi-District Comparative Design

This study compares student achievement among school districts with and without magnet-based public school choice policies. For its school choice policy information this study draws on a 1991-92 telephone interview of a nationally representative stratified random sample of 602 school districts (Steel & Levine, 1994); for its student achievement data this study uses math and reading achievement scores from the National Assessment of Educational Progress (NAEP); demographic information to adjust comparisons among the districts comes from the National Center for Education Statistic's Common Core of Data

Before proceeding to the details of the data, design, and measures, a methodological comment on selection bias is warranted. The main strength of this study, other than its large sample and national focus, is that its comparative design minimizes the form of selectivity bias that afflicts the magnet versus nonmagnet comparisons discussed earlier. As previously described, when one compares the students in a district's magnet schools to those not in the magnet schools, there are likely confounding effects of self-selection: Within a district, when parents sort themselves by choice into magnet schools and nonmagnet schools, children of the magnet school parents often have higher test scores because of family background variables. A more ideal approach would be to compare student achievement between school districts, with and one without school choice, but identical in all other ways (e.g., size, urbanicity, student demographics, educational expenditures). If magnet schools contribute to higher student achievement, either because they themselves are better schools, or because they engender competition that makes all schools better, then the district with school choice should have higher achievement. Such a difference could reasonably be taken as evidence of effects of school choice.

Comparing school districts attenuates self-selection issues. While research indicates that within a school district, magnet school parents differ from nonmagnet school parents, there is less reason to believe that at the district level there are large differences in the characteristics of parents between school districts with magnet-based school choice policies than those without them. While it is conceivable that cities with school systems with magnet-based choice policies might differentially attract or retain higher SES parents, our methodology controls for key demographic variables. Furthermore, if there is self-selection of higher-achieving students to magnet districts, even after demographic controls are applied, then one would expect higher NAEP achievement in the magnet school districts. The methodology used in this paper explicitly tests district effects and should reveal this effect if it exists.

Data, Sample, and Measures

The data for this study came from three sources:

Interview information on student assignment policies. In 1991-1992 a telephone interview was conducted with district officials in 602 school districts to collect information on magnet schools and other types of policies for school choice and desegregation in a national sample of school districts (Steele & Levine, 1994). The purpose of the telephone interview was not to elicit respondent's opinions. Rather, the purpose was to collect information on types of student assignment policies used in the districts in the sample (e.g., school choice or other types of student assignment policies), whether or not the districts had magnet schools, how many magnet schools districts had, the types of magnet schools in operation, and other district policies related to school choice (e.g., transportation support, parent information systems) and school desegregation goals. Because of the fact-gathering nature of the telephone interview, interviews were conducted with individuals in the school district office knowledgeable about student assignment policies, magnet programs, and school desegregation. Multiple interviews were sometimes conducted, with supplementary information collected from district policy documents. Letters were sent to all districts selected for the sample in advance of the interview explaining the purpose of the study and arranging the interviews. The telephone interviews were conducted by American Institutes for Research, using a computer-assisted telephone interview system.[8] Of the districts targeted in the sampling frame, 1% refused to participate, and for another 5%, no data were collected due to difficulties with scheduling or making contact with an appropriate district respondent within the time frame of the study.

A stratified sampling procedure was used to develop the district interview sample. The primary purpose was to maximize the precision of estimates of student participation in, the characteristics of, and the prevalence of magnet schools and school choice in the nation's public school systems. It was known before the design of the sample that magnet schools, school choice, and school desegregation policies are quite common in larger urban districts, fairly common in mid-size districts, and not common in tiny districts with only a few schools.[9] Therefore, to increase the precision of estimates, the sample was designed to most closely reflect this universe of school districts. The nation's 155 largest school districts were all included in the sample, districts that were overwhelmingly one race were undersampled, and school districts that did not offer the structural potential for school choice—that is, they did not have at least two schools at one grade level—were excluded from the sample. The final sample of 602 districts was drawn from a universe of about 6,400 districts based on 1990 data.[10] As described in more detail below, about half of the 602 districts were also part of the NAEP 1992 sample. Merging the two samples created the final sample for our study.

Common Core of Data. The National Center for Education Statistics maintains a data set on public education with enrollment, demographic, staffing, and financial information (school, district, and state). This data set was used for the demographic information on the schools and districts in the school choice policy interview sample.[11] To adjust our comparisons among the schools and districts for demographic characteristics, we used the school level measure of family median income, and the district level measure of the percent of children in poverty (the latter being derived from 1990 U.S. Census data).

National Assessment of Educational Progress. Student achievement was measured by student scores on the 1992 National Assessment of Educational Progress (NAEP) in 4th grade reading and 4th and 8th grade mathematics. NAEP 1992 data were used because this was when the student assignment policy information was collected. Most magnet school programs were implemented during the late 70s and 80s, spurred by the availability of federal funding[12] and growing public and political support for voluntary alternatives to mandatory reassignment policies for school desegregation. That NAEP was administered in math and reading in 1992 at which time the magnet/public school choice survey was administered is optimal for this study for it was the 1980s that brought the widespread implementation of magnet schools and their attendant school choice policies. This study assumes if district level effects of magnet schools and choice have occurred, these effects should be measurable by 1992.

The use of NAEP achievement scores for this study carries certain advantages and disadvantages. NAEP achievement measures are administered via a balanced incomplete block spiraling design. The advantage of such a design is that it allows for a large number of achievement items to be given across the test population without any given student being burdened by too large an assessment. A disadvantage for some research purposes is the fact that any given student has responded to very few items and thus estimates of ability via item response theory methodology may contain too much error to ignore. Therefore, NAEP provides five plausible value estimates for each student conditioned on two sets of variables: responses to that portion of the assessment administered to the student as well as student background characteristics. As pointed out in the 1994 NAEP Technical Report (Allen, Johnson, Mislevy, & Thomas, 1994), the plausible values are not test scores for individuals in the usual sense, but they do estimate population characteristics. In the context of statistical modeling with plausible values, special methods must be utilized in order to obtain correct estimates and standard errors. These methods are incorporated in the HLM software and were utilized in this paper.

The issue of inferences to relevant populations represents another limitation in the use of NAEP for this study. Specifically, NAEP used a complex, multi-stage probability sampling design to select public and non-public schools as well as a probability sampling of students within those schools defined by certain ages or certain grades (Allen, Johnson, Mislevy, & Thomas, 1994). Oversampling of non-public schools and schools with large minority populations was necessary for representativeness. To ensure proper inferences to relevant populations, NAEP provides sampling weights associated with school and students that relate to the units' probability of selection into the sample. However, the sampling design of NAEP does not include districts. As such, there are no weights associated with districts to allow inferences to the population of school districts in the United States, or even the largest districts in the United States. Thus, it is important to recognize that the results reported below are unweighted and do not represent inferences to the population of districts, schools, or students.

These limitations notwithstanding, we consider the use of NAEP as the measure of achievement as justified for several reasons. First, NAEP is a highly regarded test of major U.S. national policy significance. The test is a product of extensive research and development by national experts in curriculum, instruction, and measurement and employs broad participation and review in item development.[13] NAEP uses both multiple choice and free response items. Second, unlike high-stakes tests used in many states and districts, teachers have no incentive to "teach to" NAEP; schools and districts, since they are not identified, have little incentive to "look good" on NAEP. Therefore, there is no reason to suspect test score inflation biases in the sample. Third, while there is likely to be variation among the districts in the sample in the degree to which their reading and mathematics curriculum align with NAEP's content specifications, there are no a priori reasons to assume systematic differences creating biases affecting the adjusted comparisons of the models estimated in our analyses. Fourth, NAEP uses an extremely large national sample of students. No other assessment measure meeting all of the criteria above, and also providing a large sample is available to conduct the kind of policy study reported here.

Final sample. Our final sample, then, consists of a large number of racially diverse, multi-school districts, about half of which have school choice policies of some form. The sample intentionally over-represents larger districts because it is in these systems where magnet-based school choice policies are predominantly found. The districts for our final sample are those with both student assignment policy information and student/school NAEP scores. Even though this study uses just a subset of NAEP data—only those districts that were in both the 1992 NAEP sample and the 602 district school choice study sample—the merged sample still is large: more than 30,000 students, 1,000 schools, and 300 districts.

Coding of the sample. About 300 districts had both student assignment policy information (e.g., information on magnet schools and associated programs and policies) and NAEP scores, which lead to a sample of about 30,000 students in 1,000 schools. With the (district policy) interview information, it was possible to code the students and schools in the sample according to their district student assignment policies. The main question on the interview form used for this coding asked whether the district had magnet schools. Magnet schools require school choice policies. A specific definition of magnet schools was read to the respondent to insure reliable responses to this question.[14] Additional questions asked about transportation and parent information support. Other questions asked whether the district had a "majority to minority" form of choice policy—an approach that does not specifically use magnet schools, that allows parents some school choices, and that regulates enrollments in the interest of racial desegregation; and whether the district used neighborhood schools-based student assignment or rezoning and busing policies for school desegregation.

The policy information from the telephone interviews was used to code the student/school data sample and conduct several comparative analyses:

High Mag v. *Low Mag* v. *Other*. The student/school data were coded as to whether they operated in a "high mag," "low mag," or "other" (no school choice) environment. Both "high mag" and "low mag" students/schools were in a school choice policy environment, but a distinction was made so that two dichotomous variables were created, thus permitting a more robust test of the magnet/school choice variable.

The "high mag" versus "low mag" distinction was created because – as the market theory goes—the more school options are available with supporting transportation and information systems, the more the system approximates an education marketplace with presumed productivity benefits stemming from competition. In theory, school systems that provide a large supply of magnet schools and support parental/student access with well developed transportation and information systems create a more open and robust market. For instance, if a school system has 30 schools and only 2 are available as choice schools, and there is no public transportation available for these schools, and little advertising – it is hard to argue that this system resembles a market, although technically such a system has school choice. On the other, a market is produced, arguably, when there are large numbers of accessible choice schools supported by transportation and parental information systems. The interview information on district school choice systems was used to categorize systems as "High Mag" in which more than 20% of their schools were magnet schools, free bus transportation was provided to all schools in the system, and choice was supported with parental information systems using at least three different media (e.g., a district radio station, newsletters, and magnet school open houses). Districts that did not achieve all three of these characteristics were classified as "Low Mag." (Note the terms "high" and "low" are arbitrary.) The use of two dichotomous variables (rather than just one) lessens the chance that an absence of observed effects might be due to some of the students/schools operating in an environment classified as "school choice," but in reality offering relatively fewer options.

Neighborhood v. *Other*. The student/school sample was coded as to whether the policies were of the "neighborhood schools" type or "other." A "neighborhood schools" policy refers to assigning children to a school based on their residence within an attendance area boundary which is drawn to encompass the neighborhoods surrounding and closest to the school. "Other," in this classification refers to either choice policies, which could include magnet schools, or various types of mandatory reassignment for school desegregation purposes.

"Other" Choice v. *Magnets Or No School Choice*. The student/school sample was coded as to whether the policies permitted selected types of school choice through forms other than those associated with magnet schools or did not. The "other" category exists because the survey sought to identify kinds of policies that technically fall under the rubric of "choice," though in fact these are often fairly limited forms of choice, such as allowing transfers among schools through special request. Some districts have pre-existing arrangements with one or several neighboring districts and allow limited numbers of student transfers. Some districts allow transfers among schools where the transfer improves racial balance — that is, parents may select a school outside their neighborhood if the chosen school is predominantly composed of students of a different race. These modest forms of choice constitute the "other" category.

We should emphasize that the primary contrast of interest is the first described above: "high mag v. low mag v. other." This is the strongest test of the magnet-based school choice model. In the "high mag" and "low mag" districts large percentages of parents in the district have genuine choices among a variety of schools with distinctive themes and programs and the schools in these districts operate in an environment that places schools into more of a competitive relationship than they would otherwise be in a traditional "nonchoice." The "other" category, to which "high mag" and "low mag" districts are compared consists of districts in which students' placement in schools is by central administrative assignment, which does not give school choices to parents.

However, all three different types of contrasts described above have been examined. First, it was possible to do so based on the information from the telephone survey; second, these different student assignment policy options are included in discourse and debates about the pros and cons of different forms of student assignment policy, including all the different forms of public school choice. Therefore it is worth examining if they appear to be associated with any differences in achievement as measured by NAEP.

Additional sample information. Tables 1 through 4 provide information on the sample. These figures are from the telephone survey (district magnet status), from the CCD for the school year 1991-1992 (% free lunch; % black), and from 1990 U.S. census (adult education; family income). Table 1 compares the demographics of the districts with and without magnet schools. The main difference in these two categories of districts is that the districts with magnet schools (constituting approximately one third of the districts) are on average larger and have higher percentages of minority and low-income students. As other studies have shown, magnet schools are far more likely to be in urban districts with high minority populations (Blank, Dentler, Baltzell & Chabotar, 1983; Steel & Levine, 1994)—not surprising because magnet schools are associated with racial desegregation programs. The districts with magnet school/ public choice policies have on average a little more than twice as many schools and students: a mean of 79 schools in magnet districts and 36 in nonmagnet districts; and a mean of 49.5 thousand students in magnet districts versus 23 thousand in nonmagnet districts. The percentage of black students and free-lunch eligible students is higher in magnet districts by, respectively, about 10 percent and 14 percent. The adult populations living in the school districts with magnet schools are also demographically of somewhat lower socio-economic status than the populations in districts without magnet school/public choice policies.

Table 1. Demographics of "Magnet" versus "Nonmagnet" Districts

District Magnet Status	% Free Lunch	% Black	% Poverty	Adult Education	Family Income
No	32.4	14.6	17.3	12.6	40308
Yes	46.7	24.7	23.8	12.6	38403
Total	37.3	18.1	19.6	12.6	39646

Table 2. Demographics of "Neighborhood Schools" versus "Not Neighborhood Schools" Districts

Neighborhood Schools?	% Free Lunch	% Black	% Poverty	Adult Education	Family Income
No	37.3	20.7	21.6	12.6	38290
Yes	29.7	13.3	15.9	12.6	42153
Total	37.3	18.1	19.6	12.6	39646

Table 3. Demographics of "Other Choice" Districts

Other Choice?	% Free Lunch	% Black	% Poverty	Adult Education	Family Income
No	37.2	19.6	19.5	12.6	40026
Yes	36.3	14.5	19.2	12.7	38980
Total	36.9	18.2	19.4	12.7	39813

Table 4. Demographics of Parent Sample and NAEP Subsample

"Parent" Sample	% Free Lunch	% Black	% Child Poverty	Adult Education	Family Income
All District	34.5	14.8	18.5	12.7	40507
NAEP subsample	36.0	18.1	19.7	12.6	39671

Table 2 provides comparative demographic information for the districts classified as utilizing a neighborhood schools policy as compared to all other districts. Districts with neighborhood schools tend to be smaller by about 17.5 thousand students on average, and have lower percentages of black students and students eligible for free lunch, by about 8 percent and 12 percent, respectively. This, presumably, reflects the fact that larger districts with higher percentages of black students are the ones that either through court or school board directives, have developed alternative student assignment policies in an effort to reduce racial isolation in schools.

Table 3 shows relatively similar characteristics of the districts classified as "other choice" as compared with those classified as "no school choice." As described earlier, from either a theoretical or practical perspective this contrast is the least interesting of the three. Presumably because each of these two categories represents something of a mixture of student assignment policy types, there appear not to be demographic covariates of the policy categories.

As described above, the sample for this analysis is created out of the common cases from two different samples. It should be noted that while the original telephone survey sample (n=602) was a national random sample, the extent to which the sample for this analysis is truly random is unclear. There is no a priori reason to assume the sample on which our analyses are based is not random or nearly random, for it is created out of two different samples, both of which are stratified random samples. Comparing the analysis sample to the original "parent" sample of 602 reveals few noteworthy differences as shown in Table 4. The means of the demographic characteristics of the analysis sample as compared with the parent sample in most cases differ by just a few percentage points; just as important, there is no pattern to the differences, suggesting these differences where they occur are random.[15]

Analytical Methods

The same general analytic strategy is employed for each of the different types of comparisons. Because of the nested structure of the data (with students nested in schools that are, in turn, nested in districts) we employ three-level hierarchical linear modeling, using the software program HLM, to correctly account for the nested structure of the data. (Bryk, Raudenbush, & Congdon, 1994). We extend previous research by considering the effects of school choice as a three level model. By specifying a three-level model with choice modeled at the district level, we address the multilevel nature of the data in a way that has not been done before in school choice research, as well as address the issue of selection bas - arguing that selection bias is much more of a problem when choice is measured at the school level than at the district level.

Our analytic approach specifies a series of steps. To begin, we assess the proportion of variance in student achievement accounted for by differences among students, schools and by districts. This is accomplished by decomposing the total variation in student achievement across students, schools, and districts, without adding predictors at any level. Next, we add the choice comparison at the district level in a three-level model and recompute the variance accounted for, conditional on the choice comparison. Next, we add statistical controls at each level, one at a time. At the student level, we add level of parental education (PARED) as a proxy for student socio-economic status, at the school level we add school median income (SCHINC), and at the district level we add the percent of students living below poverty (POVERTY). Our interest is in determining the effect of school choice on student achievement after controlling for PARED, SCHINC, and POVERTY. Therefore, we *grand-mean center* PARED, SCHINC, and POVERTY yielding results that can be interpreted as adjusted achievement means, adjusted for the statistical controls at each level.[16] For each step we report the reduction in the variance at each level as well as changes in the mean difference in adjusted student achievement as a function of the choice policy. A more complete discussion of the statistical model employed in this study is given in Appendix 1.

Our statistical comparisons (e.g., "magnet" v. "nonmagnet") were adjusted using these demographic variables for both theoretical and practical reasons. The main theoretical reason is that, whether measured at the individual level or aggregate levels, student achievement has a strong correlation with family socio-economic background variables (Jencks, 1972; Kerckhoff, 1996). Theoretically, the effects on productivity of new policies and reform initiatives should operate independently of the demographic backgrounds of the students (this assumes that the policies/initiatives are not of a type intended specifically to be more effective for certain demographic types of students.) Theories about effects on productivity of market forces engendered by school choice do not make distinctions about differential effectiveness related to student background. Thus, it is reasonable to assume that if magnet-based school choice creates the kinds of market forces envisioned by advocates of choice, these forces should produce observable achievement differences adjusting for demographic characteristics of the student populations served.

As in all research of this type, the demographic variables we used in our statistical models were determined by what was available in the data sets. At the student level, NAEP does not have a measure of family or household income and does not collect free-lunch eligibility data. Survey information collected by NAEP asks students to report the level of academic degree of their parent(s)/guardian(s), ranging from some high school to graduate degree (PARED). At the school level, NAEP provides information based on census data on the average income of the neighborhood where the school is located (SCHINC). At the district level, we used percent of children in poverty (POVERTY) because this variable is widely accepted as an appropriate measure of the demographic factors that create challenges for school districts' education mission. When people describe the population served by a school district as "disadvantaged" or "impoverished," it is the proportion of children living in poverty they refer to, not so much parents' educational background or how many middle or upper income families there are, although it is understood that all these variables are correlated with one another.

Results

The results are organized with respect to the various school choice comparisons. For each comparison, we examine its effect on 4th grade reading achievement, 4th grade mathematics achievement, and 8th grade mathematics achievement. In what follows, Model 0 refers to the unconditional model; Model 1 adds the district choice comparison at the district level; Model 2 adds parental income at the student level; model 3 adds school median income at the school level; Model 4 adds district poverty at the district level.

Magnet v. non-Magnet Districts

Table 5 shows the result of the multilevel analysis comparing magnet districts to non-magnet districts on 4th grade NAEP reading achievement. Model 0 displays the breakdown of the percent of variance accounted for by differences among students, schools, and districts. It can be seen that only about nine percent of the variance in 4th grade reading achievement can be accounted for by differences among districts. The majority of the variance (approximately 73%) can be accounted for by differences among students, with about 19% accounted for by differences among schools. Model 1 displays the Magnet school effect without controlling for parental education, school median income, or district poverty. Here we find significant differences in favor of non-magnet districts. High magnet districts show a greater difference than low magnet districts when compared to non-magnet districts. Models 2 through 4 add parental education, school median income, and the number of children living below poverty in the district, respectively. It can be seen that for each additional control variable the differences among magnet and non-magnet districts decrease. Substantial decreases are observed when adding school median income and district level poverty. It may be interesting to note that adding parental education does not account for a substantial amount of variation in student level achievement, whereas the addition of school median income and poverty result in a comparatively larger decrease in variance explained by schools and districts respectively.

In Model 4, it can be seen that the high magnet district's coefficient remains significant, although the size of the remaining difference is very small (-5.9 points), a difference of about 1/7th of a standard deviation. Thus, the mean NAEP score of the high magnet districts is slightly lower than the mean of the comparison group, defined as districts without magnets and school choice.

Table 5. Multilevel Models of the Effect of Magnet Schools on NAEP 4th Grade Reading Achievement

	Model 0	Model 1	Model 2	Model 3	Model 4
Fixed Effects					
District Mean	210.812	208.916	208.879	208.874	212.235
Low Mag		-5.235*	-5.091*	-2.542	-0.786
High Mag		-11.239*	-11.063*	-8.856*	-5.895*
Variance Components					
Student Level	989.906	989.922	984.768	984.850	984.835
School Level	254.605	253.551	253.501	181.854	182.184
District Level	120.084	102.913	103.020	65.860	48.378
Variance Decomposition (Percentage by level)					
Student Level	72.54	73.52	73.42	79.90	81.10
School Level	18.66	18.83	18.90	14.75	14.99
District Level	8.80	7.64	7.68	5.34	3.98

Model 0: Fully unconditional model
Model 1: School Choice policy added at district level
Model 2: Model 1 + parental education added at student level
Model 3: Model 2 + school income added at school level
Model 4: Model 3 + district poverty measure added at district level

* $p < .05$

Table 6 shows the results for the multilevel analysis comparing magnet districts to non-magnet districts on 4th and 8th grade NAEP mathematics achievement. Again, it can be seen that individual differences among students account for the vast majority of the total variation in mathematics achievement scores. Difference among districts account for only a small fraction of the variance. In addition, it can be observed that the differences among magnet and non-magnet districts shrink after accounting for demographic characteristics of students, schools, and districts; only the high magnet category retains a statistically significant

Table 6. Multilevel Models of the Effect of Magnet Schools on NAEP 4th and 8th Grade Mathematics Achievement

Effect	Model 0		Model 1		Model 2		Model 3		Model 4	
	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th
Fixed Effects										
District Mean	213.526	260.368	211.930	258.093	216.425	258.093	216.156	259.587	214.746	262.239
Low Magnet			-4.982*	-7.516*	-4.961*	-7.556*	-2.642	-6.637*	-0.776	-2.826
High Magnet			-10.207*	-11.617	-10.111*	-11.688*	-8.336*	-10.461*	-4.770*	-5.231*
Variance Components										
Student Level	762.968	997.441	762.969	999.416	761.539	996.525	765.972	996.455	766.930	995.006
School Level	179.044	152.473	178.714	151.879	178.868	150.054	122.522	115.343	123.416	115.862
District Level	132.829	180.478	117.529	159.690	117.726	158.014	70.411	100.122	45.743	60.225
Variance Decomposition (Percentage by level)										
Student Level	70.98	74.97	72.03	76.23	71.97	76.38	79.88	82.22	81.93	84.96
School Level	16.66	11.46	16.87	11.59	16.90	11.50	12.78	10.00	13.18	10.00
District Level	12.36	13.57	11.10	12.18	11.12	12.11	7.34	8.26	4.89	5.14

Model 0: Fully unconditional model
Model 1: School Choice policy added at district level
Model 2: Model 1 + parental education added at student level
Model 3: Model 2 + school income added at school level
Model 4: Model 3 + district poverty measure added at district level

* $p < .05$

coefficient, although the effect size is very small—a little over 1/8th of a standard deviation.

The largest decrease in the difference among magnet and non-magnet districts occurs after adding district poverty. In terms of explained variance, it can be seen that adding school median income and district poverty respectively explain substantially more variation in math achievement compared to adding parental income at the student level.

Neighborhood v. "Other" Districts

Table 7 shows the result of the multilevel analysis comparing districts with neighborhood assignment polices to districts without neighborhood assignment polices on 4th grade NAEP reading achievement. Model 0 displays the breakdown of the percent of variance accounted for by differences among students, schools, and districts. It can be seen that only a fraction of the variance in 4th grade reading achievement can be accounted for by differences among districts. Model 1 displays the neighborhood district effect without controlling for parental education, school median income, or district poverty. Here we find significant differences in favor of neighborhood districts. Models 2 through 4 add parental education, school median income, and the number of children living below poverty in the district, respectively. It can be seen that for each additional control variable the differences among neighborhood and non-neighborhood districts decrease as does the percent of explained variance, with the most dramatic decrease occurring when adding district poverty.

Finally, Table 8 shows the results for the multilevel analysis comparing neighborhood districts to non-neighborhood districts on 4th and 8th grade NAEP mathematics achievement. As with the previous analyses, differences among districts account for only a fraction of the variance. In addition, it can be observed that the differences among neighborhood and non-neighborhood districts decrease after accounting for demographic characteristics of students, schools, and districts. In addition, as with reading, the largest decrease in the difference among neighborhood and non-neighborhood districts as well as percent of variance explained occurs after adding district poverty.

"Other" Choice v. Magnets Or No School Choice Districts

Across both grade levels and for both reading and mathematics achievement, there were no significant differences in mean district achievement scores among districts that reported having school choice and those that did not. Therefore, there was no need to control for parental education, school median income, or district level poverty.

Table 7. Multilevel Models of the Effect of Neighborhood Schools on NAEP 4th Grade Reading Achievement

	Model 0	Model 1	Model 2	Model 3	Model 4
Fixed Effects					
Average District Mean	210.771	208.916	208.879	208.874	210.510
Neighborhood Schools		6.013*	6.094*	6.093*	1.183
Variance Components					
Student Level	995.957	995.960	991.396	991.358	991.477
School Level	255.032	254.938	254.441	254.564	182.441
District Level	118.969	111.530	112.115	112.068	53.486
Variance Decomposition (Percentage by level)					
Student Level	72.70	73.10	73.01	73.00	80.78
School Level	18.62	18.71	18.74	18.75	14.86
District Level	8.68	8.48	8.26	8.25	4.17

Model 0: Fully unconditional model
Model 1: School Choice policy added at district level
Model 2: Model 1 + parental education added at student level
Model 3: Model 2 + school income added at school level
Model 4: Model 3 + district poverty measure added at district level

* $p < .05$

Adding control variables either did not change the mean differences appreciably or the mean differences became even smaller.

Discussion

Our findings are inconsistent with optimistic claims about the possibility of magnet-based school choice engendering district-wide productivity gains through forces of parent choice and school competition; our findings are more consistent with views that, whatever salutary benefits individual magnet schools may have, there is no reason to expect that systemically magnet/choice districts are very different from districts without magnet schools and with student assignment policies linked more with zoned attendance areas.

None of our statistical models produce evidence of a strong association between district-level student assignment systems and student achievement measured on NAEP in reading and mathematics. Indeed, the only statistically significant associations uncovered indicated a small difference in NAEP scores favoring the nonchoice districts compared with

Table 8. Multilevel Models of the Effect of Neighborhood Schools on NAEP 4th and 8th Grade Mathematics Achievement

Effect	Model 0		Model 1		Model 2		Model 3		Model 4	
	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th	Math 4th	Math 8th
Fixed Effects										
District Mean	213.490	260.382	211.930	258.093	211.946	258.093	213.261	259.587	224.222	274.471
Nbrhood effect			5.014*	6.803*	4.938*	6.777*	2.297	4.421*	0.226	0.912
Variance Components										
Student Level	774.946	999.774	774.944	999.777	768.947	997.315	769.025	997.266	769.025	997.278
School Level	173.757	151.642	174.119	151.298	174.042	150.923	122.650	115.803	123.537	116.918
District Level	134.751	180.531	127.960	171.495	127.596	170.686	79.099	114.037	48.127	63.161
Variance Decomposition (Percentage by level)										
Student Level	71.52	75.06	71.95	75.59	71.82	75.62	79.22	81.27	81.75	84.70
School Level	16.04	11.38	16.17	11.44	16.25	11.44	12.63	9.34	13.13	10.00
District Level	12.44	13.55	11.88	12.97	11.92	12.94	8.10	9.29	5.12	5.36

Model 0: Fully unconditional model
Model 1: School Choice policy added at district level
Model 2: Model 1 + parental education added at student level
Model 3: Model 2 + school income added at school level
Model 4: Model 3 + district poverty measure added at district level

* $p < .05$

just the high magnet districts (but no difference when compared with the magnet districts with a lower prevalence of magnet schools). We suggest caution in interpreting the substantive significance of the small, but statistically significant differences between the categories of districts for several reasons: first, the large sample size (approximately 30,000 students) allows small differences to achieve statistical significance; second, it is known from other research studies that districts implementing magnet schools have histories of racial segregation and on average are more likely to be districts encompassing large numbers of inner city schools – schools that may appear demographically similar on available measures but that in reality often face extraordinary challenges of family dislocation, student residential mobility, and neighborhood drugs and violence (Jargowsky, 1996; Wilson, 1997); and third, much research shows that magnet schools themselves, if not their districts as a whole, tend to have relatively high performance.

Below, we present first, two interpretations of these results we call, respectively, "zero-sum reallocation" and "adverse reallocation." Then, because some research indicates magnet schools often have superior academic outcomes, we offer two interpretations that may reconcile our findings with this previous research.

Zero-sum reallocation suggests that the kinds of school choice policies examined in this study essentially re-allocate students among schools, but generate no endogenous productive forces. That is, magnet schools and the school choice policies in which magnets operate may create choices and educational options for families, and even have other benefits, but they do not create market forces making schools better in ways that lead to higher academic achievement. School choice, in this view, is largely a process of redistributing students among schools within the district. There may be reasons to do this and benefits from this, but this process does not affect achievement on a district-wide level enough to be detectable by NAEP assessments of reading and mathematics when compared with districts that assign students to their schools, with little or no family choice involved. (This view is based on the cautious interpretation of our results discussed earlier that our findings do not produce evidence to infer important substantive differences between the categories of districts as measured by NAEP scores.)

Is zero-sum reallocation inconsistent with other research indicating magnet schools have higher achievement scores? It is not inconsistent if magnet schools' higher achievement scores result from self-selection processes *within* districts. Assuming a constant composition of students within the district,17 then the movement of higher-scoring students from other schools to magnets raises magnet schools' test scores at the expense of other schools. This sort of reallocation will not raise the district average. If school choice—at least the forms of public school choice examined here—does not do much more than reallocate students among schools, then there is little reason to expect average scores among school choice districts to be higher than average scores among districts without school choice policies.

A second more critical interpretation—adverse reallocation—suggests magnets may actually have adverse effects within a district. This interpretation would focus on the small negative coefficients of the high magnet category of districts. Some argue magnet schools may have a detrimental effect on nonmagnet schools by drawing off the best students, in effect depleting nonmagnets of model students that would otherwise be present to the benefit of less able students; some also contend magnets relegate nonmagnets to a "second class" status, with adverse morale consequences for staff morale and parent support (Bastian, 1985; Carrison, 1981; Moore and Davenport, 1989; Pearson, 1993).

On the other hand, if we assume that magnet schools do indeed produce higher achievement among their students and we assume that this does not come at the expense of nonmagnet schools,18 then districts with a sizable number of magnet schools should have higher average achievement than districts without magnet schools. Perhaps this is so, and NAEP achievement tests are not sensitive enough to measure differences in achievement between students in the two different policy contexts (magnets/school choice versus no magnets/no school choice). In our sample, of those districts with magnet schools, three-fourths had one-third or less of their schools as magnet schools; half had 12.5% or less of their schools as magnets. Thus, even if magnet schools' achievement is greater, at a district level the difference may be insufficient to boost NAEP reading and math means in the school choice category (relative to the means of the comparison groups) by statistically significant margins.

A second possibility is that higher achievement from magnets/school choice is in knowledge or skill areas different from those areas assessed by NAEP mathematics and reading tests. As described earlier, magnet schools in the great majority of cases have "specialty" areas—distinctive pedagogies or curricular foci (Metz, 1986; Steel & Levine, 1994). Conceivably, achievement differences distinguishing magnet schools from other schools without these specialty areas are not effectively measured with conventional standardized tests of reading and mathematics. Thus, there may indeed be differences—but NAEP does not measure them.

These alternative scenarios should guide future research. The severe paucity of research exploring links between district or state level school choice policy and academic outcomes allows a number of plausible interpretations of our findings.

It is important to properly delimit this study to forms of public school choice policy associated with magnet schools. There are a growing number of approaches to school choice, many more now than there were ten years ago. According to some, the "public school only" choice models—particularly the kind examined in this study—are not system-changing reforms (Chubb & Moe, 1990; Erickson, 1982; Everhart, 1982; Friedman & Friedman, 1981; Lieberman, 1986; Harmer, 1994, Peterson, 1990). They do not create market forces powerful enough to change the operation and effectiveness of public schools. According to this view, true education markets can only occur if there are many schools within a given area that are not managed by the traditional public education bureaucracy, these schools are accessible and known to parents, school performance is effectively measured and publicly reported, and all schools have incentives to attract and satisfy their customers and disincentives to lose or dissatisfy customers. There is growing support for tuition voucher and charter school policies because they more closely satisfy these conditions.

Finally, a methodological note is warranted concerning assumptions about the classification of the school districts in the sample based on the telephone interview data. For analytic purposes, the districts in the sample are placed in various dichotomous categories. We realize this is a simplification and recognize that school choice in practice is a complex operation and in practice districts do not fall neatly into dichotomous categories in their approaches to assigning students to schools. There are districts that would say "no, we do not have a school choice policy" that in fact may take a fairly liberal approach in allowing student requests to transfer among schools; and there are districts that would answer "yes" to the presence of magnet schools and school choice that upon close inspection may arguably have a fairly restricted approach to school choice. While there is no reason to suspect systematic misreporting among the telephone interview respondents, we would be remiss not to acknowledge the simplification that is imposed by placing districts in a few broad categories. It would be ideal to have a more comprehensive base of information on student assignment policies and practices for a large sample of districts, enabling more refined classifications along a continuum ranging from centralized assignment to progressively greater degrees of school choice.[19] This would be a productive direction for future research.

Notes

This paper was supported by Grant # R308F60010 from the U. S. Department of Education, OERI National Institute on Educational Governance, Finance, Policy-Making and Management. Opinions reflect those of the authors and do not necessarily reflect those of the granting agency.

[1] For more on these themes, see the "Magnet Schools of America" website: http://www.magnet.edu/

[2] Ibid.

[3] "Selective" means based in teacher recommendations and in some cases GPA or test score-based admissions criteria. This study used tests from the Second International Mathematics/Science Study rather than "off-the-shelf" standardized tests because they are more advanced, and were more suitable for the sample in this study, and because they are considered to be high quality academic tests. For more information see *Science Achievement in Seventeen Countries: A Preliminary Report* (New York: Pergamon Press, 1988).

[4] For instance, parents choosing magnet schools tend to be more informed about educational options in their district and have higher educational attainment. It is possible that magnet students' higher achievement owes in part or even entirely to their parents' and their own motivational or academic ability traits, and not to the differential effectiveness of magnet schools. Although some studies have employed sophisticated controls for family background and/or student prior achievement, selection bias remains a concern.

[5] Gamoran (1996) further assessed the effects of non-independence of students within schools utilizing a multilevel analysis and, in addition, examined the effect of selection bias resulting from non-random assignment to these various schools. Although there were several methodological issues that raise some cautions, the multilevel analyses further supported the positive effect of attending magnet. To control for selection, Gamoran specified a Heckman-type two-stage model where, in the first stage, a selection model was specified as a function of eighth grade achievement, sex, race/ethnicity, SES, and family composition plus a variety of measures that Gamoran argued tapped into reasons why parents might choose a magnet school unrelated to achievement advantages. In the second stage, Gamoran incorporated the probability of attending a magnet school derived from the selection model into the regression model for achievement and obtained adjusted estimates of the effect of school type on achievement, adjusted for the non-random assignment of students to schools. A limitation of Gamoran's study is that he did not incorporate the selection model into the multilevel analysis of the effects of choice on student achievement. Thus, while Gamoran correctly recognized the problems associated with clustered student data and non-random selection, he did not combine the two methods so as to control for both problems simultaneously.

[6] For a review of these and other empirical studies of school choice see Goldhaber, 1999.

[7] Howell et al. (2001) respond to these methodological issues.

[8] This system helps insure valid coding of data. The system automatically performed "allowed-value" checks and would not permit entry of out-of-range data. For example, "yes/no" items were restricted to yes or no responses (including "don't know," "refused," "inapplicable"); numeric responses were restricted to a reasonable range of values.

[9] This was known from other policy information sources (e.g., compliance records from the Office of Civil Rights as well as funding and program information managed by the federal Magnet School Assistance Program) and from previous research (e.g., Blank, Dentler, Baltzell, & Chabotar, 1983).

[10] For details of the sampling frame, see Appendix A (Steele & Levine, 1994).

[11] School-level free-lunch data for a number of school districts were missing in the CCD data set for the years we needed. For these districts, we contacted directly the individuals responsible for managing the district free-lunch data and obtained in many cases the needed school-level free-lunch counts. Data fields for districts without school-level free-lunch figures were coded as "missing" and not computed in our analyses.

[12] The Emergency School Aid Act passed in 1975 and subsequent legislation (Education for Economic Security Act, Magnet Schools Assistance Program) helped support the development and implementation of magnet school programs. For more on this, see Glenn (1978), and Warren (1978).

[13] For more details on this, see the "Data Compendiums" for NAEP 1992 reading and mathematics available from NCES.

[14] "A magnet school is defined as a public school whose primary purpose is to meet *all three* [emphasis original] of the following goals: to offer a special curricular theme or method of instruction, such as math/science, performing arts, or open classroom; to attract at least some students voluntarily from outside an assigned neighborhood attendance zone; to improve desegregation by meeting specific race/ethnic goals (i.e., it must have some type of race/ethnic controls or targets." Respondents were free to ask questions of the interviewer for additional clarification of terms.

[15] We conducted T-tests on selected sets of means and found no statistically significant differences (P=.05). Mag v Nomag on free-lunch eligible percentages & percentage black students.

[16] Multilevel modeling requires that a choice be made with regard to the centering of predictor variables. The issue of centering concerns a rescaling of predictor variables for the purposes of providing interpretable intercepts of a linear model. In standard linear regression, centering is usually not an issue because the intercepts are typically not quantities of substantive interest. In multilevel regression however, intercepts are very important quantities reflecting variability in the average values of within group variables across groups. The choice of centering determines the interpretation of the intercepts. (See Bryk & Raudenbush, 1992 pp 26-27 for details.)

[17] Some evidence, however, indicates magnets may actually retain or attract students who otherwise would not be enrolled in the district – they would otherwise be in a private school or attend a school outside the urban system in which magnets ordinarily exist.

[18] Several of the studies of magnet schools presented earlier used controls for self-selection processes and produced evidence that magnets' higher achievement is not merely a selection phenomenon, but a result of processes endogenous to magnet schools.

[19] See Archbald (1996) for more theoretical and methodological elaboration on this idea.

References

Allen, N. L., Johnson, E. G., Mislevy, R. J., and Thomas, N. (1994). Scaling Procedures. In *The 1994 NAEP Technical Report*. Washington DC: National Center for Education Statistics, U.S Department of Education.

Archbald, D. (1995). A longitudinal cohort analysis of achievement among elementary magnet students, neighborhood school students, and transfer students. *Journal of Research and Development in Education*, 28(5), Spring, 161-169, 1995.

Archbald, D. (1996). SES and demographic predictors of magnet school enrollment. *Journal of Research and Development in Education*, 23(3), Spring, 152-163.

Archbald, D. (1996). Measuring school choice using indicators. *Journal of Educational Policy*, 10 (1), March, 88-108, 1996.

Armor, D. H. (1989). After busing: Education and choice. *Public Interest*, 95, Spring, 24-37.

Barr, R. (1982) Magnet schools: an attractive alternative. *Principal*, 61(3), 37-40.

Blank, R. (1989). *Educational effects of magnet schools*. Prepared for the Conference on Choice and Control in American Education, University of Wisconsin-Madison and for the National Center of Effective Secondary Schools, University of Wisconsin-Madison.

Blank, R. & Archbald, D. (1992). Magnet schools and issues of educational quality. *Clearinghouse*, 66 (2), 81-86.

Blank, R., Dentler, R., Baltzell, C. & Chabotar, K. (1983). *Survey of magnet schools: Analyzing a model for quality integrated education*. Prepared by Abt Associates for U.S. Department of Education, Wash. DC.

Bryk, A., Raudenbush, S., & Congdon, R. (1994). HLM: Hierarchical linear and non-linear modeling with the HLM 2/L and HLM/L Programs. Chicago: Scientific Software International.

Carrison, M. (1981:34). Do magnet schools really work? *Principal*, January, 32-35.

Center for the Study of Public Policy (1970). *Education vouchers*. Cambridge, MA.

Chubb, J, & Moe, T. (1990). *Politics, markets, and schools*. Washington, DC: Brookings Institution.

Clewell, B. C. & Joy, M. F. (1993). Choice in Montclair, New Jersey: A policy information paper. In D. Waldrip, W. Markes, & N. Estes (Eds.) *Magnet school policy studies and evaluations* (pp. 58-94). University of Houston, Houston TX: International Research Institute on Educational Choice.

Clinchy, E. & Cody, E. (1978). If not public choice, then private escape. *Phi Delta Kappan*, 60(4), December, 270-273.

Coons, J. & Sugarman, S. (1977). *Education by choice*. Berkeley, CA: University of California Press.

Doyle, D. (1984). Family choice in education. Washingtion, DC: National Institute of Education.

Duke, D. (1976) Challenge to bureaucracy: the contemporary alternative school. *The Journal of Educational Thought*, 10, 34-48.

England, R. & Morgan, D. (1986). *Desegregating big city school districts: Strategies, outcomes, impacts*. Millwood, NY: National University Press.

Erickson, D. (1982). Disturbing evidence about the one best system. In Robert Everhart (Ed.) *The public school monopoly* (pp. 393-418). Cambridge, UK: Ballinger.

Everhart, R. (Ed.) (1982). *The public school monopoly*. Cambridge, UK: Ballinger.

Finn, C. (1985) Education choice: Theory, practice, and research. Testimony of Chester Finn, Assistant Secretary, Office of Educational Research and Improvement, U.S. Department of Education before the Senate Subcommittee on Intergovernmental Relations Committee on Governmental Affairs.

Finn, C. (1987) *Education Week*, 6 (39), June 24,Sec. C.

Fantini, M. (1976). *Alternative education: A sourcebook*. New York: Doubleday.

Fliegel, S. (1993). *Miracle in East Harlem: The fight for choice in public education*. New York: Times Books.

Friedman, M. (1973). The voucher idea. *New York Times*, Sep. 23.

Friedman, M. & Friedman, R. (1981). *Free to choose*. New York: Avon Books.

GAO (2002). Insufficient research to determine effectiveness of private education comparies. Report to the Ranking Minority Member, Subcommittee on the District of Columbia, Committee on Appropriations, House of Representatives (GAO-03-11). Washington DC: United States General Accounting Office.

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis*, 18(1), 1-18.

Gill, B. P., Timpane, P. M., Ross, K. E., & Brewer, D. J. (2001). Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools. Santa Monica, CA: RAND.

Glenn, J. (1978). Legislative implications of magnet schools. In Estes, N. & D. Waldrip (Eds.) *Magnet schools: Legal and political implications*. Piscataway, NJ: New Century.

Goldhaber, D. (1997). Public and private high schools: Is school choice an answer to the productivity problem? *Economic of Education Review*, 15, 93-109.

Goldhaber, D. (1999). School choice: An examination of the empirical evidence on achievement, parental decision making, and equity. *Educational Researcher*, 28 (9), 16-26.

Greeley, A. (1987) Freedom of choice: our commitment to integration. In *Parents, teachers, and children: Prospects for choice in American education* (pp. 183-207). San Francisco: Institute for Contemporary Studies.

Harmer, D. (1994). *School choice: Why you need it – How you get it*. Washington, DC: CATO Institute.

Hoxby, C. (1996). How teachers unions affect education production. *Quarterly Journal of Economics*, CXI, 671-718.

Hoxby, C. (1998). Analyzing school choice reforms that use America's traditional forms of parental choice. In P. E. Peterson & B. C. Hassel (Eds.) *Learning from school choice* (pp. 133-156). Washington DC: Brookings Institution.

Howell, W., Wolf, P., Petersen, P. & Campbell, D. (2000). Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D. C.: Evidence from Randomized Field Trials Paper prepared for the annual meetings of the American Political Science Association, Washington, DC, September 2000.

Howell, W., Wolf, P., Petersen, P. & Campbell, D. (2001). The effects of school vouchers on student achievement: A response to critics. Boston: Occasional Paper, Program on Education Policy and Governance, Harvard University.

Jargowsly, P. (1996). *Poverty and place: Ghettos, barrios, and the American city*. New York: Russell Sage Foundation.

Jencks, C. (1972). *Inequality; A reassessment of the effect of family and schooling in America.* New Yprk: Basic Books.

Jencks, C. & Areen, J. (1971). Educational vouchers: a proposal for diversity and choice. *Teachers College Record*, 72, 327-336.

Katz, E. & Danet, B. (1968) *Bureaucracy and the public*. New York: Basic Books.

Kerckhoff. A. (1996). *Generating social stratification : Toward a new research agenda*. Boulder, CO: WestviewPress.

Kolderie, T. (1985). Two alternatives routes to the improvement of education: Part II. Report of the Public Services Redesign Project. Minneapolis, MN: Humphrey Institute, University of Minnesota.

Larson, J., Witte, J., Staib, S., & Powell, M. (1990). A microscope on magnet schools: Secondary school magnet programs. Montgomery County Public Schools; Rockville, MD: Carver Educational Services Center.

Levine, D. & Havighurst, R. (1977). *The future of big city schools: Desegregation policies and magnet alternatives*. Berkeley, CA: McCutchan.

Levine, D. (1975). Home environment of students in a high achieving inner city parochial school and a nearby public school. *Sociology of Education*, 45, Fall, 435-445.

Lieberman, M. (1986). *Beyond public education*. New York: Praeger.

Lin, Q. (2002). An evaluation of charter school effectiveness. *Education*, 122 (1), Fall,166-177.

Lipsky, M. (1980) *Street level bureaucracy*. New York: Russel Sage Foundation. Manley-Casimir, M.(Ed.) (1982) Family choice in schooling. Lexington, MA: Lexington Books.

Manley-Casimir, M.(Ed.) (1982). *Family choice in schooling*. Lexington, MA: Lexington Books.

Metz, M. (1986). *Different by design*. New York: Routledge & Kegan Paul.

Michaelson, J. (1981) A theory of decision making in the public schools: a public choice approach. In S. Bacharach (Ed.) *Organizational behavior in schools and school districts*, (pp. 208-244). New York: Praeger.

Miron, G., & Nelson, C. (2001). Student academic achievement in charter schools: What we know and why we know so little. New York: National Center for the Study of Privatization in Education.

Moe, T. (2001). Teachers unions and the public school. In T. Moe (Ed.) *A primer on America's schools* (pp. 151-184). Stanford University, Stanford CA: Hoover Institution Press.

Nathan, J. (Ed.) (1989). *Public schools by choice: Expanding opportunities for parents, students, and teachers*. Bloomington, IN: Meyer-Stone.

NCES. (1988). *National Educational Longitudinal Study of 1988*. Washington, DC: U.S. Department of Education.

Osborne, D. (1999). Healthy competition. *The New Republic*, October, 31-35.

Peterson, P. (1990). Monopoly and competition in American education." In Clune, W. and J. Witte (Eds.) *Choice and control in American education* (pp. 47-79). New York: The Falmer Press.

Peterson, P. & Hassel, B. C. (Eds.) (1998). *Learning from school choice.* Washington, DC: Brookings Institution.

Peterson, P. (2001). Choice in American education. In T. Moe (Ed.) *A primer on America's schools* (pp. 249-284). Stanford University, Stanford CA: Hoover Institution Press.

Powell, A. (1996). *Lessons from privilege: The American prep school tradition*. Cambridge, MA: Harvard University Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Second ed.). Thousands Oaks, CA: Sage Publications.

Raywid, M. (1985) Family choice arrangements in public education. *Review of Educational Research*, 55(4), Winter, 435-467.

Rossell, C. (1990). *The carrot or the stick for school desegregation policy?* Philadelphia, PA: Temple University Press.

Smrekar, C. & Goldring, E. (1999). *School choice in urban America: Magnet schools and the pursuit of equity*. New York: Teachers College Press, Columbia University Press.

Steel, L. & Levine, R. (1994). *Educational innovation in multiracial contexts* (Prepared by American Institutes for Research). Washington, DC: U.S. Department of Education.

Teelken, C. (1999). School choice in the Netherlands, England and Scotland in a comparative perspective. *Comparative Education*, 35, 283–302.

Toch, T., Linnon, N., & Cooper, M. (1991, May 27). Making kids smarter: How magnet education is giving children the edge they need (cover story). *US News & World Report*, pp. 58-66.

Van Zanten, A. (1996.) Market trends in the French school system: overt policy, hidden strategies, actual changes. *Oxford Studies in Comparative Education*, 6, 63-75.

Young, T. & Clinchy, E. (1992). *Choice in public education*. Columbia University, New York: Teachers' College Press.

Waldrip, D., Markes, W., & Estes, N. (1993). *Magnet school policy studies and evaluations*. University of Houston, Houston TX: International Research Institute on Educational Choice.

Warren, C. (1978*). The magnet school boom: Implications for desegregation*. New York, Columbia University. ERIC No. ED 157 965.

Wells, A. (1991). *The sociology of school choice: A study of black students' participation in a voluntary transfer plan*. Unpublished doctoral dissertation. Columbia University.

Wilson, W. J. (1997). *When work disappears: The world of the new urban poor*. New York: Vintage Books.

Whitty, G., Power, S. & Halpin, D. (1998). *Devolution and choice in education: The school, the state and the market.* Buckingham, UK: Open University Press.

Appendix 1

We employ a three-level hierarchical linear model with randomly varying intercepts. For the purposes of this study, interest centered on modeling changes in average achievement (as estimated by the intercept of the model) as a function of school choice indicators after controlling for student, school, and district characteristics. The relationships between these different demographic characteristics and achievement are not the focus of this study. Therefore, the flexibility of hierarchical linear modeling allows us to focus on changes in the intercept while allowing slope variability without attempting to model that variability.

We tested the normality of the distributions of NAEP's plausible value scores for 4th grade reading and math (1992) and for 8th grade math (1994). In each instance the distributions were highly normal as evidence by skewness scores averaging -0.19 and kurtosis scores avering -0.10.

The specification of the three-level model is as follows. The level 1 model can be written as

Level-1: $$Ach_{ijk} \quad {}_{0jk} \quad {}_{1jk}(PARED)_{ijk} \quad e_{ijk}, \qquad \text{(A1)}$$

where Ach_{ijk} is the relevant NAEP achievement score for student *i*, in school *j*, in district *k* and *PARED* is the measure of parental education for student *i*, in school *j*, in district *k*. Assuming the *PARED* is grand-mean centered, p_{0jk} is the adjusted mean achievement scores that are assumed to vary across schools and across districts, p_{1jk} is the slope representing the relationship between student achievement and student *PARED*, and e_{ijk} is a random error term.

The Level-2 model can be written as

Level-2: $${}_{0jk} \quad {}_{0k} \quad {}_{1k}(SCHINC)_{jk} \quad r_{jk}, \qquad \text{(A2)}$$

where b_{0k} is the adjusted school mean achievement scores, adjusted for the percent of students who are free lunch eligible, b_{1k} is the slope representing the relationship between school mean achievement school median income, and r_{jk} is a random error term.

The Level-3 model can be written as

Level-3:

$${}_{0k} \quad {}_{000} \quad {}_{001}(POVERTY)_k \quad {}_{002}(CHOICE)_k \quad u_{0k}, \qquad \text{(A3)}$$

where g_{000} is the grand-mean achievement score, g_{001} is the slope representing the relationship between district level poverty and district level student achievement, g_{002} is the slope representing the school choice comparison of interest, and u_{0k} is a random error term. In the level-3 model, grand mean centering is not employed for *CHOICE*. Estimation of the model parameters utilizes full-information maximum likelihood estimation of the variance components along with empirical Bayes estimation of the fixed effects of the model (Raudenbush and Bryk, 2002).