On School Choice and Test-Based Accountability

Damian W. Betebenner
Boston College

Kenneth R. Howe
Samara S. Foster
University of Colorado


Abstract
Among the two most prominent school reform measures currently being implemented in The United States are school choice and test-based accountability. Until recently, the two policy initiatives remained relatively distinct from one another. With the passage of the No Child Left Behind Act of 2001 (NCLB), a mutualism between choice and accountability emerged whereby school choice complements test-based accountability. In the first portion of this study we present a conceptual overview of school choice and test-based accountability and explicate connections between the two that are explicit in reform implementations like NCLB or implicit within the market-based reform literature in which school choice and test-based accountability reside. In the second portion we scrutinize the connections, in particular, between school choice and test-based accountability using a large western school district with a popular choice system in place. Data from three sources are combined to explore the ways in which school choice and test-based accountability draw on each other: state assessment data of children in the district, school choice data for every participating student in the district choice
program, and a parental survey of both participants and non-participants of choice asking their attitudes concerning the use of school report cards in the district. Results suggest that choice is of benefit academically to only the lowest achieving students, choice participation is not uniform across different ethnic groups in the district, and parents’ primary motivations as reported on a survey for participation in choice are not due to test scores, though this is not consistent with choice preferences among parents in the district. As such, our results generally confirm the hypotheses of choice critics more so than advocates. Keywords: school choice; accountability; student testing.

Introduction

During the last two decades, The United States has witnessed a sweeping tide of reform efforts directed toward the improvement of public education. Largely initiated by the seminal report A Nation at Risk (National Commission on Excellence in Education [NCEE], 1983), the quest for improved education in the United States ventured in numerous directions—from the varied curriculum wars of the 1980s to the standards based reform efforts of the 1990s. With reauthorization of the Elementary and Secondary Education Act of 1965 (ESEA), referred to as the No Child Left Behind Act (NCLB), no abatement is apparent in the desire to reform education at the beginning of the 21st century. Two reform initiatives currently dominate public education policy: Test-based accountability and parental choice. Given their inclusion in NCLB, both are likely to pervade discussions concerning education reform for years to come and to have lasting effects on public education in The United States.

NCLB is without precedent in both scope and direction, employing test-based accountability and school choice as the fundamental mechanisms to engender school improvement. The legislation implements this accountability-choice mutualism along two lines: First, NCLB encourages the expansion of school choice by allowing students attending Title I schools to transfer to another public school (including charter schools) if their school has been identified as being in need of school improvement, corrective action, or restructuring as defined by the accountability components of the law. NCLB provides technical assistance for schools failing to make adequate progress and also requires school districts to provide free transportation for students who choose to attend other district schools. Second, NCLB requires statewide assessments in grades 3–8 in reading and math, and science testing in at least one grade in elementary, middle, and high school. Furthermore, beginning in the 2002–03 school-year, states were required to provide report cards for each of their public schools.

In simplest terms, the rationale for combining choice and accountability is that the two complement one another: (1) accountability facilitates parental choice by supplying parents with information (in the form of school report cards) they need to make informed choices, and (2) choice facilitates accountability by fostering competition for enrollment. With the choice/accountability nexus as the foundation, our intent in this study is twofold: First, we wish to analyze the theory that brings accountability and choice together under one federal policy. In particular we wish to unpack the rationale linking choice and accountability using the language of market-based reforms and make explicit the manner in which accountability and parental choice are thought to complement one another. Second, using data compiled from a large western school district, we ascertain whether market-based reforms that link accountability and choice function as prescribed. We begin with a brief review of market-based education reforms vis-à-vis school choice.
Market-based School Reforms

Milton Friedman (1962) is widely credited with originating the idea of reforming public education via school choice. He proposed a market-driven scheme in which public funding and administrative authority would be transferred to private schools where parents redeemed their government supplied vouchers. Various other voucher proposals have followed that have added the overall improvement of public education to Friedman’s more modest goal of increased efficiency (e.g., Chubb & Moe, 1990). A few limited voucher programs have been implemented, with Milwaukee’s “Parental Choice Program” being the most long-lived and carefully studied (see Witte, 2001). The Florida A+ program (a hybrid system, which includes both private and public schools) has gained considerable attention of late (Greene, 2001). Since vouchers were declared constitutional by the United States Supreme Court in 2002, a number of states are in the process of considering legislation that would open the door to vouchers in their states. Colorado, for example, recently passed a state voucher program that allocates funding to disadvantaged children from kindergarten through grade 12 whenever their local school district demonstrates low academic performance. The program was recently struck down by the Colorado State Supreme Court. But proponents intend to reintroduce a voucher bill that will pass constitutional muster. It’s likely that other states will follow Colorado’s lead in using choice as a major initiative to reform public education.

Not all proponents of public school choice embrace a market rationale, but this rationale now sets the terms of the debate (Hess, 2002). Proponents argue that market regulation is far better than bureaucratic government regulation (Chubb & Moe, 1990). The market mechanism of parental choice allows funds for schools to follow students and creates a system of competition among schools. They argue that this competition provides schools with a mandate to improve or risk losing money and even closure (Greene, Peterson, & Du, 1997; Hess, Maranto, & Milliman, 1999; Hess & Leal, 2001; Merrifield, 2001). The result is an incentive for schools to provide better services (through, for example, innovation, specialization, efficiency, etc.) and to increase student achievement. Other supporters contend that such programs can serve to promote equity (Coons & Sugarman, 1978). In general, those who see public choice as a means of promoting equity observe that parents have long chosen schools by choosing their place of residence (Henig & Sugarman, 1999). Parents’ incomes and social positions thus largely determine their power to choose. A choice policy that removes attendance boundaries permits students to choose schools independent of the price of houses in the neighborhoods in which they live. It thus provides all parents with choice, and also promises to promote diversity in schools

Critics of market-driven school choice question whether it can improve achievement overall. They contend that the market may simply redistribute students as a result of “skimming,” where certain schools’ mean achievement increases only because other schools’ mean achievement decreases (e.g., Carnoy, 2000). Critics also contend that by introducing competition, cooperation that currently exists between teachers and schools will be sacrificed. The majority of criticism for market-based reforms like school choice falls on the issue of equity, charging that such reforms are much more likely to exacerbate inequity than to mitigate it. School choice plans, for example, typically make no provision to protect students from being harmed while schools are declining, before they are reconstituted (e.g., Arsen, Plank & Sykes, 2000; Lauder & Hughes, 1999). School choice can also result in stratification by ethnicity and income (e.g., Cobb & Glass, 1999; Howe, Eisenhart & Betebenner, 2001), exclude special needs students (e.g., Arsen et al., 2000; Rothstein, 1999; Zollers, 2000), and thereby force other public schools to carry the burden of accommodating the needs of more difficult to teach students. Finally, critics also claim that there is inequality among parents in their capacity to choose because certain parents may lack the information needed to participate in
meaningful deliberation, and others may lack trust in authorities (Fuller & Elmore, 1996; Wells, 1993).

In a comprehensive overview of the effects of school competition and educational outcomes, Belfield and Levin (2002) conducted a review of research from the United States in order to determine “to what extent, and according to what measures of output, does increased competition improve educational quality” (2002, pp. 279–280). The criteria for the articles chosen were that they must address educational outcomes and competitive pressures across large markets and contain an analysis on the basis of large-scale, cross-sectional data sets. Competition was constructed in terms of greater school choice for parents and students, and educational outcomes were mostly measured in terms of standardized test scores, as well as graduation/attainment, expenditures/efficiency, teacher quality, students’ post-school wages, and local housing prices. They identified about 40 relevant empirical studies and determined that the research “shows reasonably consistent evidence of a link between competition (choice) and educational quality” (p. 297). In other words, they found a positive correlation between increased competition and educational outcomes/quality, but the actual effects were “substantively modest” (p. 297). They raise caution in using these findings to support policies to promote greater competition among schools. They suggest that the benefits of increased competition must be set against any additional generated costs to justify specific policy approaches. They write, “The benefits of competition...should not be exaggerated. ... a number of them may in fact be the ‘same’ benefit, but calculated in different ways” (p. 296). Additionally, they warn that equity issues stemming from increased competition must be considered since “market systems rank poorly against equity criteria (e.g., by showing greater segregation and partitioning of student groups)” (Belfield & Levin, 2002, p. 296–297).

**School Choice and Test-based Accountability**

Despite their marriage in NCLB, school choice and test-based accountability were not originally envisioned as complementary policy initiatives. Their paths toward policy prominence have been different and often uneven. Though the roots of both reside in the 1960s, it was test-based accountability that first gained a foothold under Title I of the Elementary and Secondary Education Act of 1965. With its implementation more than three decades ago, large scale testing has been the primary means of evaluating program efficacy and the determination of federal support for the education of low achieving students in poor neighborhoods (Heubert & Hauser, 1999). Since then, student testing for accountability purposes has only grown in prominence.

The mid-1980s marked an expansion in the use of students’ standardized test results for accountability purposes (Linn, 2000, 2001; National Research Council (NRC), 1999). In 1986, 33 states required some form of minimum competency testing of its students (Office of Technology Assessment, 1992). Following the Charlottesville Summit in 1989, state and national leaders launched a more ambitious conception of accountability based upon standards-based reforms. The tests used for this new system of accountability were to be aligned to world-class standards and were intended to be more intellectually challenging than were those of the 1980s. As the 1990s unfolded, standards-based accountability became the “touchstone” for state governance, as states moved away from judging schools in terms of inputs (Elmore, Ablemann, & Fuhrman, 1996) and focused on results. By the mid-1990s, for example, test based requirements for high school graduation existed in 18 states (Bond, Braskamp, & Roeber, 1996).

Toward the latter half of the 1990s, a subtle yet important shift in test-based accountability occurred: the results of assessments began to be distributed as broadly as possible to the general public. The primary vehicle for providing this information has been “school report cards”. By 2001,
27 states had implemented some form of school report cards as part of their accountability systems (Orlofsky & Olson, 2001), and, as mentioned previously, report cards are mandated by NCLB. Test results are the primary ingredient of the report cards, which also include information such as safety, class demographics, graduation and dropout rates, student mobility, and information on student teacher ratios (Goertz, Duffy, & Carlson-LeFloch, 2000; Education Commission of the States, [ECS], 1999). Today, school report cards are available online, from the district, and are often published in local papers.

The rationale underlying support for test-based accountability—whether in the 1960s or currently—is that it enables school improvement by tying rewards and sanctions to measurable outcomes as defined by standardized test scores. Whereas school administrators have a great deal of power to alter schools based upon test results, parents, regardless of what information is available to them, have been largely powerless to implement school level change. Thus, in order to extend the rationale underlying support for test-based accountability to situations where parents become the primary consumers of the testing results, parents need to have at their disposal some means by which to effect change upon the school. Giving parents the choice to determine what school their child attends is one such means.

Some early studies (e.g., Wells, 1993; Wells & Crane, 1992) suggest that only savvy and powerful parents avail themselves of accountability information, and that they use it to promote only their own children’s interests. More recent studies (e.g., Schneider, Teske & Marshal, 2000) concur that only a relatively small group of activist parents use such information, but go onto suggest that the behavior of these parents (even if self-interested) produces overall improvement. To date, no studies have been conducted on the universal school report card requirements of the type required by NCLB—in which all parents receive accountability information, not just those who actively seek it out—to address fundamental questions such as: What kinds of parents use the information provided, for what purposes, and to what effects?

The debate over public school choice is currently driven by a collection of isolated and conflicting studies. In their recent comprehensive review of the existing empirical evidence, Gill, Timpane, Ross, and Brewer (2001) assert that no general conclusions are warranted with regard to the efficacy of school choice programs. The authors cite two major shortcomings with regard to existing research on school choice: (1) Current research is of the “black box” variety—not delving deeply into how choice affects the actual operation of the school and merely focusing on inputs and outputs, and (2) current research has focused on programs that have been operating for a relatively short period of time with few participants. Another shortcoming is that the conflicting studies have examined different educational policy contexts or, in the case of Cobb and Glass (1999) versus Hoxby (2001), used different data and analysis methods to examine the same context.

The most contentious issue for school choice policies historically—and that now extends to policies such as NCLB that integrate choice and accountability—is their effect on patterns of school enrollments by race/ethnicity, income, and special education. Some scholars have argued on behalf of choice—for vouchers targeted at low-income students, in particular—that it provides a means by which to increase racial integration (e.g., Greene, 1999), but this claim has been disputed (e.g., Reardon & Yun, 2002). The effects of choice on patterns of enrollment, it would seem, are highly dependent on the details of specific policies and the social and historical context in which they operate (e.g., Gorard, et al. 2001, Wells, 1998). Several studies of comprehensive choice programs (i.e., programs not confined to targeted vouchers) have supported the more modest claim that while choice may not increase integration neither does it increase stratification (e.g., Gorard, et al.; Hoxby, 2001). This conclusion has been disputed, however, by studies where choice has increased stratification (e.g., Carnoy, 2000; Cobb & Glass, 1999; Fiske & Ladd, 2000; Howe, et al. 2001).
Another family of studies of the effects of choice has focused on the academic achievement of students enrolled in choice schools (e.g., charter schools and means-tested vouchers). No general conclusions have emerged regarding whether choice schools outperform assigned public schools with comparable students. Moreover, these studies are limited because they fail to examine the effects of choice policy on students in the system remaining in assigned neighborhood schools. A comprehensive evaluation must also include the effect of choice on achievement (positive or negative) of students remaining in assigned schools (e.g., Gill, et al. 2001; Hoxby, 2001).

Several recent studies have adopted this more comprehensive approach. For example, Gorard, Fitz, and Taylor (2001), Greene (2001), and Hoxby (2001) have each produced empirical studies supporting the argument that competitive pressures created by choice policies (in the form of charter schools and/or vouchers) foster improved achievement overall, inclusive of low performing schools. Although these studies are an advance over the kinds of studies described previously, they have important limitations with respect to the kinds of policies examined vis-à-vis NCLB-like policies. In particular, each was based on school or district level data (as opposed to student level data) with only the Gorard et al. study employing longitudinal data.

The Current Study

Following recent studies, the goal of this study is to examine school choice and test-based accountability together in an empirical fashion so as to provide a more comprehensive report on each and a more insightful synthesis of market-based reforms in general. To this end our study addresses questions in three areas:

School choice and student achievement. An outcome suggested by much of the school choice literature is that students participating in a choice program will outperform their counterparts who do not. Our study investigates this question and also includes the effect of school choice on achievement of students not participating in choice and remaining in their neighborhood schools.

School choice and patterns of student enrollment. Another set of analyses in this study focus on the nature of the students and schools participating in school choice. Analyses here fall into two broad categories: (1) what are the characteristics of students participating in school choice, and (2) what are the characteristics of the schools that are most desirable in terms of choice?

Parents’ use of accountability information. Data on achievement and the movements of students within a district allowing for choice is crucial to understanding how achievement and choice interact. But it is the dissemination of achievement data to parents that is at the heart of the school choice/test-based accountability mutualism. To include this component in our study we surveyed parents both participating and not participating in choice and asked them about their use of school report cards and other information with regard to beliefs about their children’s schools.

Method

The location for the present study is a large western school district with an approximate enrollment of 27,500 students. Historically, the quality of the district’s schools has ranged from very good to excellent. School choice has existed in the district since 1961. The choice program in the district, called open enrollment, allows parents from both inside and outside the district to send their child to any school in the district in which there is space available after enrollment by neighborhood children. By the mid 1990s, spurred by parents who were unhappy with the district’s implementation of the “middle school philosophy” or who complained about a perceived lack of emphasis on
academics in the district more generally, various choice options began to proliferate. Coincident with these developments, a new school board sympathetic to choice was elected, and the superintendent responsible for the middle school philosophy resigned. This was also a time when the school choice movement began to accelerate at both the state and national levels.

As open enrollment expanded in the district, four choice options were added to the traditional option of enrolling in any neighborhood school on a space-available basis: (1) focus schools which offer a particular curricular focus; (2) neighborhood focus schools which offer the standard district curriculum; (3) strand schools which offer the standard district curriculum alongside a different curricular strand; and (4) charter schools, whose accountability to the district is specified by a contract. In 1999–2000, 21 of 57 district schools had incorporated one of the types of choice options just described.1 This compares with only five such schools providing such choice in 1994–95, all emphasizing either bilingual or experiential education. Thus, between 1994–95 and 1999–2000, 16 additional articulated choice schools were added, half of which adopted the mission of an explicit emphasis on academic rigor and college preparation. Core Knowledge was most prominent among the new options provided with five schools adopting it. Currently, more than 25 percent of students now take advantage of open enrollment to attend the district schools other than those assigned to them by attendance area.

Concurrent with the boom in school choice in the district the state began a statewide testing program. Implemented in 1997, the program tested 4th grade students in reading and writing. The program has since expanded and currently tests students from 3rd to 10th grade in reading and writing, from 5th to 10th grade in mathematics, and science in the 8th grade. For the past three years, annual test results as well as numerous other data associated with each school are tabulated by the state department of education and reported to the public as school accountability reports (SAR). These reports are available to the public from a number of different sources including local newspapers and both the district and state education websites.

Whereas the effects of school choice are typically hard to isolate, the district considered here is a relatively closed system where schools must compete for enrollment from the same pool of students. Thus, the district provides an ideal setting to scrutinize the broad assertions leveled by proponents and skeptics of school choice in particular and market-based reforms in general. Even so, there are a number of confounding factors to be considered before any actual analysis is performed. The ideal “experimental” situation would provide for a pretest, a treatment consisting of the parents choosing (or not choosing) which school their child will attend, and post-tests to see what the results of choosing are. The two primary components of such an analysis are student assessment and parental choice data. Limitations in the availability and nature of such data require that only a small subset of all students in the district be used in the analysis of school choice and academic achievement.

Based upon previous school choice research in the district (Howe et al., 2001), we identify two groups of parents participating in choice in the district: Parents whose children are exercising choice between school level and parents whose children are exercising choice within school level. Children participating in choice between school levels are children who are enrolling in either kindergarten, 6th grade, or 9th grade and thus are, overall, transitioning into a new elementary, middle, or high schools.2 Choice exercised within school level occurs when children switch elementary, middle or high schools in mid-stream. These two groups of parents indicated fundamentally

---

1 One of two K–8 schools, 11 of 33 elementary schools, 5 of 13 middle schools, and 4 of 9 high schools.

2 The schools in the district under study are primarily K–5 elementary, 6–8 middle, and 9–12 high schools. There are two schools designated as K–8.
different motivations for participating in the district choice program. Within school level parents often exercised choice in the hopes of alleviating specific difficulties that their child was having at their current school. Thus, the rationale for participating in choice for these parents was highly idiosyncratic. By contrast, between school levels parents, particularly at the kindergarten and 6th grade levels, had more homogeneous concerns centered around their child attending the highest quality elementary or middle school available. In addition, the great majority of choice participants, 75 percent, were between school levels.

Among the between school level parents exercising choice, further limitations exist. There are numerous parents who enroll their child in kindergarten but those children are not subject to pre-testing and thus do not allow for any analysis to be performed concerning the results of choice. The ideal group of children for analysis is the group of children making the transition from elementary to middle school in the district: Academic measures exist which allow for a pre-post comparison, they are numerous, and their parents motivations for participating in choice (i.e., wanting their child to attend the best middle school possible) are relatively homogeneous. Unless otherwise indicated, the analyses performed here concern students matriculating from the 5th grade in 2000–01 to the 6th grade in 2001–02. Data for the study came from three sources: district assessment data, district choice records, and a parent survey of beliefs about school report cards:

Assessment data. Test data for all students in the district from the inception of the statewide testing program were provided by the district. Assessment data were pared to only include 5th and 6th grade students in the district during the 2000–01 and 2001–02 academic years, respectively. Tests in reading and mathematics were given to students in those grades and in those years.

Choice data. School choice records were provided by the district for analysis. These records include data regarding 3 ranked school choices, as well as the neighborhood school the student would normally attend. After substantial data cleaning, these records were combined with the test data when a matching student record could be found. After matching, 1,961 cases remained for analysis.

Survey data. A telephone survey was administered in spring 2003 to elicit parental beliefs about the school report cards. Two populations of parents were identified: The population of parents sending their child to their neighborhood school (non-choosers) and the population of parents participating in the choice program. A stratified random sample was constructed using location within district and whether the parent’s child was entering kindergarten, sixth, or ninth grade. The total number of chooser and non-chooser respondents was 200 and 202 respectively. Because survey respondents were guaranteed anonymity, no linking of parent to student data was possible.

The merged assessment and choice data was used to examine questions regarding students achievement as well as examining patterns of choice based enrollment. Survey data was used to assess parent attitudes towards the school report cards in the district. The methodology employed to address each of the three areas follows.

School Choice and Student Achievement

To quantify the academic value of choice with respect to student achievement, we employed a series of multilevel growth models (Goldstein, 2003; Singer & Willett, 2003). In general, the model employs three levels: Occasion k (level 1) within Student j (level 2) within School k (level 3). Within this three level structure a series of four nested models were analyzed: (1) an unconditional means model; (2) an unconditional growth model with no fixed effects; (3) an unconditional growth model taking account of whether or not the student is a participant in choice or not; and (4) an
unconditional growth model taking account of both student participation in choice and initial academic ability. Models were examined for both reading and mathematics.

The unconditional means model, sometimes referred to as the empty model, partitions variance between the three levels of the model and is employed as a baseline with which to compare later more complicated models. The model is given in composite form in Equation 1. The equation partitions the math/reading scale score into school \((ν_{00k})\), student \((ξ_{0jk})\), and occasion \((ε_{ijk})\) components.

\[
Y_{ijk} = γ_{000} + ν_{00k} + ξ_{0jk} + ε_{ijk}
\] (1)

The unconditional growth model fits a linear trajectory to each student in the data set. Because there are only two years of data currently available, there will be no residual error at level one due to a perfect fit. Two unconditional growth models are of interest: A growth model which sets growth to differ between students but is fixed between schools and a growth model allowing growth to vary between both students and schools. The purpose of examining the two models is to determine whether the more parsimonious model which fixes growth across schools is sufficient to model the data. Equation 2 presents the unconditional growth model allowing for slopes to vary across individuals but fixes growth across schools. Equation 3 allows growth to vary across both individuals and schools. The residuals \(ν_{00k}, ξ_{0jk},\) and \(ε_{ijk}\), are identical to those defined in the unconditional means model of Equation 1. The residuals \(ξ_{1jk}\) and \(ν_{1k}\) are residuals for slope at levels 2 and 3, respectively.

\[
Y_{ijk} = γ_{000} + γ_{100}(GRADE_{ijk} − 5) + ν_{00k} + ξ_{0jk} + ε_{ijk} + ξ_{1jk}(GRADE_{ijk} − 5)
\] (2)

\[
Y_{ijk} = γ_{000} + γ_{100}(GRADE_{ijk} − 5) + ν_{00k} + ξ_{0jk} + ε_{ijk} + ν_{1k}(GRADE_{ijk} − 5) + ξ_{1jk}(GRADE_{ijk} − 5)
\] (3)

To model the effects of choice on achievement, a fixed effect is added to the unconditional growth model. The fixed effect has three levels depending upon whether the student did not participate in school choice, participated in school choice and accepted their enrollment at a school of their choosing, or participated in school choice but attended their assigned neighborhood school. Based upon the results of Model 2 versus Model 3, growth is allowed to vary across both individuals and schools. We consider two models in this context: A parsimonious model accounting for choice that sets all choice groups’ growth rates to be equal and a model that allows for growth rates to vary between choice groups. Equation 4 provides the equation used to estimate the choice effects while fixing growth rates between groups. Equation 5 is the equation allowing for different growth rates for the three choice groups.

\[
Y_{ijk} = γ_{000} + γ_{010}CHOICE_j + γ_{100}(GRADE_{ijk} − 5) + ν_{00k} + ξ_{0jk} + ε_{ijk} + ν_{1k}(GRADE_{ijk} − 5) + ξ_{1jk}(GRADE_{ijk} − 5)
\] (4)

\[
Y_{ijk} = γ_{000} + γ_{010}CHOICE_j + γ_{100}(GRADE_{ijk} − 5) + γ_{110}(GRADE_{ijk} − 5)CHOICE_j + ν_{00k} + ξ_{0jk} + ε_{ijk} + ν_{1k}(GRADE_{ijk} − 5) + ξ_{1jk}(GRADE_{ijk} − 5)
\] (5)
The final multilevel model investigating the effects of choice on academic achievement incorporates a further fixed effect indicating the quartile of the student on the initial measure. The purpose of this is to control for initial ability since choosers are generally higher performing students in this district than non-choosers and regression toward the mean is likely to affect the attribution of growth rates to the choice groups. The fully crossed model allowing for initial status quartile and choice group is given in Equation 6.

\[ Y_{ijk} = \gamma_{000} + \gamma_{010} \text{CHOICE}_j + \gamma_{020} \text{QUARTILE}_j + \gamma_{021} \text{QUARTILE}_j \cdot \text{CHOICE}_j + \gamma_{100} (\text{GRADE}_{ijk} - 5) + \gamma_{110} (\text{GRADE}_{ijk} - 5) \text{CHOICE}_j + \gamma_{120} (\text{GRADE}_{ijk} - 5) \text{QUARTILE}_j + \gamma_{121} (\text{GRADE}_{ijk} - 5) \text{QUARTILE}_j \cdot \text{CHOICE}_j + \gamma_{021} + \nu_{00k} + \xi_{0jk} + \epsilon_{ijk} + \nu_{1k} (\text{GRADE}_{ijk} - 5) + \xi_{1jk} (\text{GRADE}_{ijk} - 5) \] (6)

**School Choice and Patterns of Student Enrollment**

Analyses on choice based student enrollment fall into two categories: First, we investigate what factors are most likely to predict whether a student is a participant in choice or not. Next, we investigate what are the characteristics of schools that draw people via the choice program. The purpose is to look at patterns of enrollment due to choice both in terms of the students and in terms of the schools. To analyze choice patterns associated with students, we employ a two-level logistic regression model using chooser/non-chooser as the dichotomous outcome variable (Snijders & Bosker, 1999). In this setting, level 1 represents the student and level 2 represents the 5th grade school the student was attending when they applied to be part of the choice program. One of the benefits of employing the two level design is that it allows for the modeling of between school variation which is known to exist (i.e., students are much more likely participate in choice in some elementary schools than in others). Like with the analysis of achievement and choice, a series of models is used to isolate the most prominent factors associated with choice participation.

We begin with the unconditional means model (or empty model). This model, given in Equation 7, expresses the log-odds of the probability of being a chooser in school \( j \), \( \logit(P_j) \), as a linear function of the grand mean of all such probabilities, \( \gamma_0 \), plus random deviations from this average for each group, \( U_{0j} \).

\[ \logit(P_j) = \gamma_0 + U_{0j} \] (7)

Next, we wish to test whether the likelihood of participating in choice is dependant upon the academic performance of the student measured the year prior to when the choice decision would take effect. When there are different likelihoods for high achievers versus low achievers, we refer to this as process as “skimming”. To test whether “skimming” is occurring with regard to test performance, 5th grade reading and math scores were added as covariates to Model 7. This model is given in Equation 8.

\[ \logit(P_{ij}) = \gamma_0 + \gamma_1 \text{READ}_{ij} + \gamma_2 \text{MATH}_{ij} + \xi_{ij} \] (8)

Lastly, we extended the model of Equation 8 to include a fixed factor for ethnicity. This inclusion allows one to test whether or not “skimming” is occurring with regard to the ethnicity. Because there are so few African American and Native American students in the district, we restrict
reporting, but not analysis, to Asian, Hispanic, and White students. This model is given by Equation 9. The estimates of the model indicate whether the likelihood of participating in choice is dependent upon the ethnicity of the student.

\[
\logit(P_{ij}) = \gamma_0 + \gamma_1 \text{READ}_{ij} + \gamma_2 \text{MATH}_{ij} + \gamma_3 \text{ETHNICITY}_{ij} + \xi_{ij} \tag{9}
\]

Results

Results from an examination of school choice and student achievement were interesting. The baseline results for Model 1 are presented in Table 1. Here again, level-one represents occasion, level-two represents students, and level three represents schools. As one would expect, most of the variability (66.4% in math and 66.9% in reading) in the observed scores is at the student level. At the school level in math and reading the amount of variability accounted for is 13.3% and 13.7%, respectively.

Extending Model 1 to include a covariate for time (Models 2 and 3), not surprisingly, provided a much better fit for the data. Model 3, which allows for differing growth rates between both students and schools provided a significantly better fit than Model 2, which allowed for differing growth rates between students and fixed growth rates for schools. In particular, for Model 3, -2LogLikelihood for Math and Reading were 40,799.720 and 38,965.620, respectively, whereas the values derived for Model 2 (which estimates two fewer parameters) were 40,815.540 and 38,982.190, respectively.

Table 1

Parameter estimates for Model 1, math and reading exams

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\gamma_{000})</td>
<td>547.0</td>
<td>645.9</td>
</tr>
<tr>
<td><strong>Random Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level-three (school)</td>
<td>703.9</td>
<td>451.1</td>
</tr>
<tr>
<td>(\text{Var. Comp.})</td>
<td>248.9</td>
<td>159.1</td>
</tr>
<tr>
<td>Level-two (student)</td>
<td>3525.9</td>
<td>2197.9</td>
</tr>
<tr>
<td>(\text{Var. Comp.})</td>
<td>133.1</td>
<td>82.3</td>
</tr>
<tr>
<td>Level-one (occasion)</td>
<td>1078.5</td>
<td>633.5</td>
</tr>
<tr>
<td>(\text{Var. Comp.})</td>
<td>35.2</td>
<td>20.7</td>
</tr>
<tr>
<td><strong>-2LogLikelihood</strong></td>
<td>41319.0</td>
<td>39353.6</td>
</tr>
</tbody>
</table>

Models 4 and 5 add a fixed effect for choice and differ with regard to their allowance for variable growth rates between three different choice groups (chooser who left (CL), choosers who stayed (CS), and non-choosers (NC)). The results are given in Table 2. The most important point to be derived from the results presented is that Model 5 was not significantly better at fitting the data than Model 5 (the difference in -2LogLikelihood between the two models for math and reading were 1.44 and 0.05, respectively) indicating that allowing equal growth rates of the three choice groups is not a hypothesis that can be rejected. This is evident given the very slight change in the variance components between the models within each subject. That is, the addition of choice status does not
help account for any of the residual variance associated with the model at across the student or school level. Overall, the results of Models 4 and 5 together confirm that there is no benefit to the academic performance of students that can be associated with their choice status in the district in the first year.

Table 2

Parameter estimates for Models 4 and 5 for math and reading exams

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Math</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 4</td>
<td>Model 5</td>
</tr>
<tr>
<td></td>
<td>Coeff.</td>
<td>S.E.</td>
</tr>
<tr>
<td>$\gamma_{000}$</td>
<td>532.7</td>
<td>6.7</td>
</tr>
<tr>
<td>$\gamma_{100}$</td>
<td>22.3</td>
<td>1.7</td>
</tr>
<tr>
<td>$\gamma_{010}$ CL</td>
<td>7.7</td>
<td>4.5</td>
</tr>
<tr>
<td>$\gamma_{010}$ CS</td>
<td>9.5</td>
<td>9.8</td>
</tr>
<tr>
<td>$\gamma_{110}$ CLxGrade</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{110}$ CSxGrade</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effects</th>
<th>Math</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level-3 school effects</td>
<td>Var.</td>
<td>S.E.</td>
</tr>
<tr>
<td>var($\nu_{00k}$)</td>
<td>733.4</td>
<td>259.5</td>
</tr>
<tr>
<td>var($\nu_{1k}$)</td>
<td>34.5</td>
<td>17.3</td>
</tr>
<tr>
<td>cov($\nu_{00k}$, $\nu_{1k}$)</td>
<td>-88.4</td>
<td>52.7</td>
</tr>
<tr>
<td>Level-2 student effects</td>
<td>Var.</td>
<td>S.E.</td>
</tr>
<tr>
<td>var($\zeta_{0jk}$)</td>
<td>4273.8</td>
<td>138.9</td>
</tr>
<tr>
<td>var($\zeta_{1jk}$)</td>
<td>1623.8</td>
<td>53.3</td>
</tr>
<tr>
<td>cov($\zeta_{00k}$, $\zeta_{1jk}$)</td>
<td>-589.3</td>
<td>62.5</td>
</tr>
<tr>
<td>Level-1 occasion effects</td>
<td>Var.</td>
<td>S.E.</td>
</tr>
<tr>
<td>var($\epsilon_{ijk}$)</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>-2LogLikelihood</td>
<td>40796.0</td>
<td>40794.6</td>
</tr>
</tbody>
</table>

Building on Model 4 and 5, the final model used to examine school choice and student achievement, Model 6, incorporates a fixed effect denoting the quartile of academic performance the student reached in the 5th grade, prior to their choice/non-choice. Including this term allows for an examination of whether choice helps/hinders students at some performance levels that are masked by the overall results of Model 3. Results of this model for both math and reading are presented in Figure 1.
On School Choice and Test-Based Accountability

Figure 1: Grade 5 to grade 6 math and reading growth estimates for non-choosers (NC), choosers leaving (CL), and choosers staying (CS) by quartile

The “gold-standard” comparison is between the choosers who left (CL) and the choosers who stayed (CS) since both groups are equivalent in the fact that they requested to leave their neighborhood schools. Because the number of CS students is small (most students that participate in choice get some school option which they take), the differences in growth rates between the CS and CL groups, though large in some cases, are not statistically significant. Comparisons between non-choosers (NC) and CL yielded interesting and statistically significant results. In particular, with regard to the math test, those students in the 1st quartile who chose out of their neighborhood school demonstrated significantly better performance on the sixth grade math test than did their counterparts who stayed. In the 2nd and 4th quartiles this ordering was reversed: students choosing out of their neighborhood schools who left performed significantly worse than did their counterparts who stayed. The results in reading demonstrated less difference between the three groups. There were no significant differences between the growth rates of the three groups in each of the four quartiles. Overall, the results suggest that no broad claims can be made about the efficacy of choice on student academic achievement as measured by the state assessment.

Results of the examination of patterns of student enrollment with regard to choice status indicate “skimming” is occurring. The logistic regression model of Equation 8 yields coefficients that were significant and positive for reading but not significantly different than zero for math. Specifically, \( \gamma_1 = 0.00103 \) (0.00117) and \( \gamma_2 = 0.00517 \) (0.00146). Thus, the probability of participating in choice increase as the student’s fifth grade reading score increases. In a choice system with no “skimming”, one would expect to see no relationship between student ability and their participation in choice. It is unclear why reading yielded a significant coefficient but not math. Certainly more data is necessary to draw long term conclusions about skimming, a likely result is movement towards higher concentrations of high ability students at certain schools and higher concentration of low ability students at other schools—tracking-at-large so to speak.
When looking at patterns of student enrollment including ethnicity as a factor some interesting results arise. Figure 2 presents the results of the analyses associated with Equation 9. Across all score levels, Asian students demonstrated a significantly higher probability of participating in choice than did Hispanic or White students. There was virtually no difference between White and Hispanic students with regard to their probability of participation. The results imply, at least in terms of those people choosing to participate in the choice system, that “skimming” based upon ethnicity is not occurring in the district given the relatively small number of Asian students in the sample.

Just as the previous results present a complicated picture of what the consequences of choice in the district are, results from the survey administered to parents in the district also indicate that the motivations behind the parents’ decision to participate or not in school choice are nuanced. Specifically we found that parents, when asked about the importance of assessment based school ratings in the decision of where to send their child, did not value the school-level test results as highly as other factors including safety and school curriculum. Parents participating in school choice did value school report cards and the test based information significantly more than did those parents not participating in school choice, but for both groups, there were other non-testing factors that were reported to be more important in their decision making process.

Figure 2: Probability of participating in choice based upon 5th grade reading scores and ethnicity

This is partly consistent with results derived from district choice data showing what schools are most popular. When looking at middle schools, where test data is available and parental influence is greatest on the school the student attends, the most requested middle schools for choice are those schools with the highest report card ratings. Thus, it appears based upon both survey data and patterns of choice in the district, that test scores do play a significant role in the decision making
process of parents. But, there are other considerations that trump test score results in the parents’ overall decision.

Conclusion

In this study we have examined arguments concerning school choice and its relation to test-based accountability in three ways: By examining how choice affects student achievement as measured by statewide achievement tests in math and reading, by examining how choice affects patterns of student enrollment across schools in the district as measured by statewide assessments, and by examining how choice is informed by results of statewide assessment that are reported in school report cards.

With regard to the assertions put forward by the proponents and opponents of choice, taking the district we examined as a crucible of choice this paper lends greater support to the contentions of opponents of choice than the proponents:

There was not a uniform benefit on student test scores for those students participating in choice. Thus, the contention that allowing choice will help academic achievement is not supported by our findings. In fact, the only students in our study who showed a positive benefit from choice were those students from the lowest quartile. But their increase was only demonstrated on the math test. It was not confirmed by their performance on the reading test.

Our results suggest that “skimming” by ability is occurring in the district. This result may go some way to explaining why, when simple descriptive statistics are analyzed at, it appears that schools with a high amount of choice do well. These schools may be equal in quality to other, less desirous schools in the district, but the pool of students from which they draw is highly able, leading some to wrongly believe that it is the school that is responsible for these children’s scores.

Parents’ decisions about what schools to send their children to are not uniformly directed by test scores alone. This is true for both parents participating in choice and those who don’t. There are many factors that influence parental attitudes about schools including reputation, safety, location, school where the child’s friends will attend, and curriculum.

In future work we hope to extend the current analyses to include more years of data that will shed light on what the longer term effects of choice are. The examination of choice and its impact upon achievement and patterns of student enrollment should be central to any discussion about the efficacy of choice in the public education system. It is our hope that similar examinations to ours will be carried out in others districts with school choice systems in place so that a more complete picture is available of this burgeoning phenomenon. In that way we all can better understand whether school choice does truly improve education in the United States.

References


About the Authors

Damian W. Betebenner  
Boston College

Kenneth R. Howe  
University of Colorado at Boulder

Samara S. Foster  
University of Colorado at Boulder

Email: damian.betebenner@bc.edu

**Damian W. Betebenner** is assistant professor in the Department of Educational Research, Measurement, and Evaluation. His research interests include longitudinal data analysis and applied statistics—particularly as they pertain to education policy.

**Kenneth R. Howe** is professor in the Educational Foundations, Policy, and Practice program area. He specializes in education policy, professional ethics, and philosophy of education.

**Samara S. Foster** is a doctoral student in the School of Education at the University of Colorado at Boulder. She is in the Research and Evaluation Methodology and Educational Foundations, Policy, and Practice programs. Her research interests include philosophy of education, philosophy of social science research, education policy, and feminist theory.
Archivos Analíticos de Políticas Educativas

Associate Editors
Gustavo E. Fischman & Pablo Gentili
Arizona State University & Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)
Roberto Rodríguez Gómez

Editorial Board

Hugo Aboites
Universidad Autónoma Metropolitana-Xochimilco

Dalila Andrade de Oliveira
Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

Alejandro Canales
Universidad Nacional Autónoma de México

Erwin Epstein
Loyola University, Chicago, Illinois

Rollin Kent
Universidad Autónoma de Puebla. Puebla, México

Daniel C. Levy
University at Albany, SUNY, Albany, New York

Maria Loreto Egaña
Programa Interdisciplinario de Investigación en Educación

Grover Pango
Foro Latinoamericano de Políticas Educativas, Perú

Angel Ignacio Pérez Gómez
Universidad de Málaga

Diana Rhoten
Social Science Research Council, New York, New York

Susan Street
Centro de Investigaciones y Estudios Superiores en Antropología Social Occidente, Guadalajara, México

Antonio Teodoro
Universidade Lusófona Lisboa,

Adrián Acosta
Universidad de Guadalajara México

Alejandra Birgin
Ministerio de Educación, Argentina

Ursula Casanova
Arizona State University, Tempe, Arizona

Mariano Fernández
Enguita Universidad de Salamanca, España

Walter Kohan
Universidade Estadual do Rio de Janeiro, Brasil

Nilma Limó Gomes
Universidade Federal de Minas Gerais, Belo Horizonte

Mariano Narodowski
Universidad Torcuato Di Tella, Argentina

Vanilda Paiva
Universidade Estadual Do Rio De Janeiro, Brasil

Mónica Pini
Universidad Nacional de San Martín, Argentina

José Gimeno Sacristán
Universidad de Valencia, España

Nelly P. Stromquist
University of Southern California, Los Angeles, California

Carlos A. Torres
UCLA

Claudio Almonacid Avila
Universidad Metropolitana de Ciencias de la Educación, Chile

Teresa Bracho
Centro de Investigación y Docencia Económica-CIDE

Sigfredo Chiroque
Instituto de Pedagogía Popular, Perú

Gaudencio Frigotto
Universidade Estadual do Rio de Janeiro, Brasil

Roberto Leher
Universidade Estadual do Rio de Janeiro, Brasil

Pia Lindquist Wong
California State University, Sacramento, California

Miguel Pereira
Catedrático Universidad de Granada, España

Romualdo Portella do Oliveira
Universidade de São Paulo

Daniel Schugurensky
Ontario Institute for Studies in Education, Canada

Daniel Suárez
Laboratorio de Políticas Públicas-Universidad de Buenos Aires, Argentina

Jurjo Torres Santomé
Universidad de la Coruña, España