Journal of Research in Rural Education, Spring, 2003, Vol. 18, No. 1, 17-30

Disentangling the Effects of Nonmetro Location and Student Poverty on School Performance/Improvement: Implications for Equitable Excellence in Kentucky Public Schools

Edward B. Reeves
Morehead State University

Kentucky public schools face a daunting challenge in meeting the state's proficiency standard by 2014. The present study evaluates the effects of student poverty and location on school accountability scores. Long-term as well as short-term models of these effects are examined using a sample of 1,115 schools nested in 171 school districts. Short-term effects are estimated with hierarchical linear modeling. The HLM results are then used to estimate how schools with hypothetical characteristics will be performing in 2014.

The HLM analysis finds that student poverty has a negative effect on the baseline school performance, and this effect is greater than that of location. The effects of location, on the other hand, are most apparent in the annual change in scores. A positive effect is associated with most nonmetro locations. These findings disclose a rural school advantage that does not lie in small schools and small districts alone. The long-term projections show poverty and location each influencing the outcomes, although the positive effect of nonmetro location does not erase the strongly negative effect of poverty. If the projections are realized, equitable excellence will not be achieved in Kentucky public schools during the foreseeable future. The paper concludes by discussing remedies for inequity and policy options for warding off the projected outcome of large numbers of public schools failing to achieve proficiency.

Background and Objectives

After "A Nation at Risk" found American schools to be so mediocre as to endanger the country's economic future, raising the level of achievement in public schools became a national *cause célèbre*. Equity has not received the same attention as achievement (Gamoran, 2000; Kahlenberg, 2000), although "equitable excellence" in the public schools should be the goal of a democratic society (Roeder, 1999). As the effort to make schools accountable for raising achievement gains momentum, continuous monitoring of school performance and improvement brings equity into the foreground of policy concerns. A concern for equity can focus on a variety of issues. In the present study, I deal with two: location and poverty.

Location versus poverty. A substantial literature is devoted to showing that rural schools are worth preserving and that, other things being equal, rural communities can provide a quality education (see Khattri, Riley, & Kane, 1997). However, as Khattri et al. report, much of the research on which the contention is based has been plagued by faulty design. Citing Coladarci (Pelavin Research Institute, 1996), they state that "the research on rural education often does not include adequate control variables, making it difficult to determine whether a particular phenomenon is truly 'rural,' or whether it is merely observed in a rural setting and could be associated with other conditions" (p. 81). Furthermore, "studies often do not include comparison groups, making it difficult to determine whether the results apply only to a rural area, or whether the results are true for other locations as well" (p. 81).

One consequence of the deficiencies noted by Khattri et al. (1997) is that previous research has often blurred the effects of poverty and location:

> Most of the data available on student outcomes are not disaggregated by location *and* by poverty, and little available research uses both variables simultaneously in examining such differences. Some of the better outcomes [that researchers have noted] for rural students (not controlling for poverty) may be a function of a smaller proportion of rural students being enrolled in high-poverty schools, as compared with students in urban areas. Thus, the degree to which student outcomes-

An earlier version of this paper was presented at the 2002 meeting of the American Educational Research Association, New Orleans.

Correspondence concerning this article should be addressed to Edward B. Reeves, UPO Box 785, Morehead State University, Morehead, KY 40351. (e.reeves@morehead-st.edu)

> for students in poor, rural areas are similar to, or differ from, those for their urban peers is not clear, and the degree to which geographic location plays a role, after poverty is taken into account, is not apparent. (p. 85)

Responding to this critique the present study employs school- and district-level data that encompass a spectrum of rural-to-metropolitan locations, a measure of student poverty at the school level, and control variables in order to disentangle the effects of poverty and location.

Excellence and equity. During the 1980s and 1990s, state governments in Kentucky, Texas, Minnesota and elsewhere implemented school assessment as part of a broad strategy to assure taxpayers and legislators that public education was held accountable for children's learning. More recently, the Federal government has weighed in on this matter, culminating with the *No Child Left Behind Act of 2001* (NCLB). The implementation of school assessment and accountability systems is now mandated for kindergarten-through-grade-12 schools in every state. Fundamental to this federal policy thrust is the insistence that states develop high standards for what students should know and learn. With standards in place, states must test every student's progress toward these standards by using tests that are aligned with the standards. Each state, school district, and school will be expected to make adequate yearly progress toward meeting the standards. School and district performance will be publicly reported. The district or school that repeatedly fails to make adequate progress toward the standards will be held accountable and will receive sanctions of various kinds.

The mandate for high standards and accountability, as embodied in state-level reform movements and more recently in the NCLB Act, can be seen as a broad policy innovation that has evolved over the course of several decades. The impetus for reform dates back to the 1954 U.S. Supreme Court decision, *Brown vs. Topeka Board of Education*, and the Coleman Report (Coleman et al., 1966). As policymakers and school systems have ventured down the path of reform, a spate of thorny issues arose with regard to the design of school assessment and accountability systems (see Ladd, 1996, National Research Council, 2001). Equity is one of the most difficult problems to confront standards-based reform to date (Gamoran, 2000). Rhetoric about the need for equitable outcomes among children of different backgrounds abounds, but the concrete means to achieve equity typically have not been undertaken by the states. Although the NCLB Act provides specific remedies for inequity, it is not clear at this time that these remedies will have the desired effect.

Equity implies that the difference in achievement between schools over time will shrink, rather than grow or remain constant. Equity is necessarily a long-term issue. No one believes that when schools are performing at very different levels, the differences can be erased overnight. Nevertheless, data collected for a period of a few years may portend a long-term trend with large implications for equity. The goal of assessing equity is better served when school performance and improvement measures actually reflect how well schools are achieving, net of student background and extrinsic differences between schools. A school whose students have higher test scores is not necessarily better than one whose students have lower tests scores. Such differences may not reflect intrinsic quality differences between the two schools, but rather may result from the different entry-level characteristics of the students or from other extrinsic factors, such as the geographic location of the school and its students. When extrinsic factors such as location and student poverty influence accountability results, the use of raw scores for accountability may be unfair and inequitable. By contrast, an accountability system may be deemed fair and equitable if it does not favor schools with students from an advantaged family background or from a particular type of geographic location. Phrased differently, the accountability system is fair and equitable if it differentiates schools based on factors that school personnel can reasonably be expected to control; fair and equitable accountability does not differentiate schools by factors that the staff cannot control (American Educational Research Association, 2000; Ladd, 1996).

The present study evaluates how much student poverty and location affect school accountability in Kentucky—in the long term as well as the short term. This objective goes beyond previous studies that have examined the effect of poverty but not location (Clotfelter & Ladd, 1996; Roeder, 1999, 2000a, 2000b) or that have assessed the short-term, but not the long-term, implications (Reeves, 2000). The present study also differs from Roeder (2001) by revealing how school location and student poverty level may have long-term consequences for whether or not Kentucky schools achieve proficiency.

Education Reform and School Accountability in Kentucky

Kentucky was in the vanguard of states when its legislature enacted the comprehensive *Kentucky Educational Reform Act of 1990* (KERA) (Elmore, Abelmann, & Fuhrman, 1996; Gamoran, 2000; Guskey, 1994). This legislation provided for the annual testing of selected grade levels within all Kentucky public schools. From the outset the testing protocol was quite innovative, even daring. It called for heavy reliance upon open-response (essay-like) questions, writing portfolios, and "performance events"—i.e., group problem solving exercises. The performance events were criticized for being difficult to score and of questionable reliability. As a result, this part of the test pro-

tocol was abandoned after a few years. Minor changes in the testing strategy were made every year with the result that comparing test scores for a period of more than a few consecutive years became problematic. Eventually, a panel of experts was invited to evaluate the assessment system. The experts raised questions about the validity and reliability of the test results (Catterall, Mehrens, Ryan, Flores, & Rubin, 1998).

Responding to the criticisms, the Kentucky General Assembly directed the Kentucky Board of Education to redesign the state's assessment system. The National Technical Advisory Panel for Assessment and Accountability, a group of testing experts from around the nation, advised the Kentucky Board of Education in its development of a new test and suggested using a model similar to that used by the American College Testing (ACT) program to predict student success in college. The new Commonwealth Accountability Testing System (CATS) was first administered in the spring of 1999. In CATS, changes were made to improve the reliability and validity of the test, to reduce the testing time, and to make the system fairer and easier to understand. Open-response questions and writing portfolios were kept as important components of the test, but nationally norm referenced (NAEP-like) questions were also included. Four years of test data using the new protocol are currently available and provide the substance for the present study.

CATS also differs from the previous Kentucky assessment tests in that it is intended to provide reliable and valid scores that can be used for student as well as school assessment. At the present time, however, only school assessment has been implemented while a study is on-going to determine how to implement the assessment of individual students. In addition to the norm referenced tests which are administered to third, sixth, and ninth graders, students in specified grades in each school are given tests that consist of multiple-choice and open-response questions in the following subject areas: reading, math, science, social studies, arts and humanities, and practical living and vocational studies. Writing portfolios and writing-on-demand exercises are given to fourth, eighth, and twelfth graders. Besides these academic indices, a nonacademic index is also created to reflect attendance and retention rates for all schools, dropout rates for middle and high schools, and successful transition to adult life for high school graduates. For purposes of assessing the performance and progress of a school, then, the scores from these various tests and the nonacademic index are averaged, weighted, and summed to give a composite Accountability Index score for each school. The highest composite score that a school can achieve is 140 points. The target mandated by the Kentucky General Assembly is for every school to reach the goal of 100 points on its composite score—a goal that is equated with proficiency—by the year 2014.[1]

Is Kentucky's reform initiative succeeding? Comparative, nationally normed data suggests that Kentucky's program has indeed been successful. The NAEP achievement gains in reading and mathematics of Kentucky schools have outstripped the national averages, and Kentucky's gains are comparable to those achieved by Texas and Minnesota, where standards-based reform was implemented earlier, in the 1980s (Gamoran, 2000). One cannot tell what the implications of these gains are for equity, however. The gains in test score levels would mean that inequality is declining only if weaker students and disadvantaged schools are posting higher gains than stronger students and advantaged schools. The state has not been quick to address this matter in a comprehensive manner, preferring instead to make note of the fact that some schools with high percentages of impoverished students are listed each year among the top performers. This is seen as evidence that disadvantaged schools should be capable of substantial progress. Recent efforts by the educational research community in the state have begun to compile a tool kit of best practices for teaching disadvantaged students, but one would have to say that this effort is still a work in progress. More recently still, in response to requirements set forth by the NCLB Act, the state has initiated a comprehensive strategy to begin collecting data on the "achievement gap" between different categories of students (e.g., minority and nonminority students, students who are eligible for free and reduced-price lunch and those who are not eligible, students with and without disabilities, etc.) (Christie, 2002).

All said, Kentucky's effort at standards-based reform has much to be admired. Along with a few other states, Kentucky is a true leader in the national standards-based reform movement. Its policymakers have followed a stable, long-term agenda avoiding the wrenching start-stop-restart pattern that has characterized reform in some other states. Reform in Kentucky was developed carefully to ensure the alignment between standards and curriculum. In the mid-1990s Kentucky was one of the few states that implemented professional development of teachers geared toward standards reform. The state has also been highly successful in garnering support from business leaders to promote reform goals and efforts (Gamoran, 2000).

Theoretical Framework

The negative effect of student poverty on between-school performance as measured by achievement tests has been repeatedly substantiated by research (for example, see Friedman & Lichter, 1998; Khattri et al., 1997; Roscigno & Crowley, 2001; Stallmann & Johnson, 1996). The link

[1]Detailed information on the assessment and accountability system may be found at the Kentucky Department of Education website: http://www.kde.state.ky.us/.

between poverty and lowered performance has specifically been noted in Kentucky (Reeves, 2000; Roeder, 1999, 2000a, 2000b).[2] Finding an association between poverty and performance does not necessarily mean that within-school gains in accountability scores are linked to poverty. However, Reeves (2000) discovered that the school district median household income explained 3% of the variance in short-term change in accountability scores in Kentucky. Roeder (2001) suggests that there is also a long-term effect of poverty on school performance, but this has not been confirmed in Kentucky prior to the present study.

Little consensus has emerged concerning the effects of geographic location on performance and improvement. In a study of NAEP data from 1975 to 1992, Greenberg and Teixeira (1995) find that the achievement gap between metro and nonmetro students narrowed dramatically during the 18-year period. Fan and Chen (1999) control for family SES and demonstrate that there are hardly any rural/urban differences in achievement on verbal, math, science, and social studies standardized tests for a national student cohort that was tested during the 8th, 10th, and 12th grades.

Khattri et al. (1997) propose numerous reasons why rural schools might not differ from metropolitan schools when poverty differences are held constant. Some researchers suggest that, other things being equal, rural schools may actually do a better job of educating than metropolitan schools. For example, Bickel, Howley, Williams, and Glascock (2001), Friedkin and Necochea (1988), Howley (1996), and Huang and Howley (1993), show that small district size and/or small school membership—both more characteristic of rural than of metropolitan school systems—have positive influences on achievement. Others report that the social capital resources of rural communities yield positive academic outcomes (Coleman, 1990; Israel, Beaulieu, & Hartlin, 2001). Contrary to these findings, Roscigno and Crowley (2001), conclude that rural location has a negative, but largely indirect effect upon academic achievement, because rural schools and rural families tend to be poorer than the national average. Thus, they support the idea that poverty and rural location are inextricably linked. Friedman and Lichter (1998), Stallmann and Johnson (1996), and Young (1998) reach similar conclusions.

Research Questions

The present study seeks empirically supported answers to three questions:

1. Does location affect school accountability results independently of student poverty?
2. Is school improvement as registered by accountability results affected differentially by school location and student poverty?
3. Having found answers to the questions above, what are the implications for school accountability, in the long term as well as the short term, and for standards-based reform?

Method

I used 2-level hierarchical linear modeling (HLM) to differentiate school-level and district-level influences on school performance and improvement respectively.

To disentangle the effects of school location and student poverty, the analysis specified student poverty as a school-level variable and location as a district-level variable. Selected control variables were also included in the HLM model (see the Appendix for a summary of the HLM models in equation format). To improve the estimation of intercepts and certain predictor variables, random effects were specified along with fixed effects (cf. Raudenbush & Bryk, 2002). The continuous predictor variables were centered on their group means if school-level, and on their grand means if district-level.

Before conducting the HLM analyses, two methodological issues were dealt with: I inspected the distributions of the independent variables to see if they looked normal. As described below, some of the variables were log-transformed in order to achieve a closer approximation to a normal distribution. The final methodological consideration before actually performing the HLM analyses was to check the independent variables for multicollinearity following a two-step process (Lewis-Beck, 1980). First, I examined bivariate correlations among the independent variables, looking for coefficients around 0.8, or higher. There were none. Since it is possible to have high multicollinearity even when the bivariate coefficients are small due to the linear combination of variables, the second step was to regress each continuous independent variable on the remaining independent variables and to look for an R^2 from each of these equations that was near 1.0. The largest R^2 found in this set of equations was 0.52. Consequently, multicollinearity was deemed to be unproblematic.

The Sample

Kentucky has 1,256 public schools and 176 school districts. One hundred and twenty of the school districts are "county" districts; 56 are "independent." Independent districts are enclaves smaller than a county. Twenty of the

[2]For a more optimistic assessment, see Weston (2000) who argues that, although poverty "sets the odds" for success in Kentucky, "schools can beat the odds" (p. 1).

Table 1
Descriptive Statistics

Variable	*M*	*SD*	Min.	Max.
School-level (N_1 = 1115)				
School performance (1999-2000 AI score)	65.77	9.72	37.45	100.90
School improvement (across two biennia)	4.24	4.60	-15.20	28.55
% students receiving free/reduced lunch	51.35	21.61	0.67	99.06
% Black students (ln)	-0.49	3.72	-6.91	4.29
School membership (ln)	6.09	0.56	4.17	7.63
Student/teacher ratio	20.43	5.20	9.40	42.10
School Type				
Elementary (reference category)	0.56	0.50	0	1
Middle	0.17	0.38	0	1
High	0.17	0.38	0	1
Combined	0.10	0.29	0	1
District-level (N_2 = 171)				
Independent district	0.30	0.46	0	1
District size (ln)	7.73	0.91	5.27	11.47
Per student revenue ($)	6459.72	655.75	5407.29	9878.58
Location				
Metro (reference category)	0.25	0.44	0	1
Nonmetro, adjacent to metro	0.24	0.43	0	1
Nonmetro, town <2,500	0.18	0.38	0	1
Nonmetro, town 2,500-9,999	0.22	0.42	0	1
Nonmetro, town ≥10,000	0.11	0.32	0	1

independent districts are located inside metropolitan counties; the rest are located in nonmetro counties. With respect to accountability, independent school districts have often performed better than county school districts (Reeves, 2000). The present study was able to make use of a sample consisting of 1,115 Kentucky public schools nested in 171 school districts.

Data and Variables

Since this study employs 2-level HLM analysis, the variables were assembled in two data sets (see Table 1).

School-level variables. To measure a baseline of school performance I used the mean Accountability Index (AI) score for the biennium of 1999-2000 school years. The 2-year mean AI score was used to reduce the volatility of single-year scores that results from testing different cohorts of students and other one-time influences (Linn, Baker, & Betebenner, 2002). Table 1 reveals that the baseline school performance score averaged 65.8. School improvement was indicated by the gain in the AI scores, between the baseline biennium (1999-2000) and the succeeding biennium (2001-2002). This calculation followed the procedure used by the Kentucky Department of Education (KDE) to determine school progress. Table 1 shows a mean 2-year improvement of 4.24 points—2.12 points when converted to an annual gain.

It is important for the long-term analysis described below to note that this estimate of annual improvement is similar to the estimate that Roeder (2001) arrived at independently. Using accountability scores for all Kentucky schools between 1993 and 2001, Roeder employed three

different methods of estimation and arrived at a best estimate of the average yearly change. His estimate of the annual change in scores was 2.34, just 0.22 higher than my estimate. This means, of course, that my projections of school performance, shown below, are slightly more conservative than his.[3]

Both of the dependent variables described above were taken directly from the accountability results reported by the Kentucky Department of Education (KDE). The remaining school-level variables included test variables and controls, the data for which were also obtained from KDE. Student poverty was measured by the percentage of students in the school participating in the free and reduced lunch program (1999). Three other variables—percent Black students, school membership, and student/teacher ratio (also dated to the 1999 school year)—were selected as control variables, because each could confound the effects of location (cf. Khattri et al., 1997). In Kentucky, a higher percentage of Black students in the school is more characteristic of metropolitan locations, as are school membership and student/teacher ratio. To correct for deviations from normality, percent Black students and school membership were normalized with natural log transformations.

Exploratory analysis of the data had revealed that the type of school had a strong influence on school performance and improvement. Thus, the effect of school type was controlled by dummy variables for middle school, high school, and combined school, with elementary school serving as the reference category. Another finding of the exploratory analysis was that the improvement in scores from the baseline biennium to the subsequent biennium was inversely related to the baseline score. This effect was controlled in the improvement model.

District-level variables. District-level variables consisted in part of the per student revenue (1999) and the natural log of district size (1999). These measures were chosen to control district-level effects that have geographically skewed distributions in Kentucky. The data for these variables were obtained from KDE. Location was another variable that was measured at the district level. Four dummy variables captured the demographic/locational features of the county where each school district was located. The dummy variables denoted different types of nonmetro school districts. They were: adjacent to metro area; town < 2,500; town 2,500 - 9,999; and town ≥ 10,000. The reference category was "metro," meaning that the school district was located in a metropolitan county. This classification scheme for location is a modification of the Urban Influence Codes obtained from the Economic Research Service, U.S. Department of Agriculture. The Urban Influence Codes have two categories of metropolitan counties (i.e., large and small). In the present study the two metro categories were collapsed into one category. The Urban Influence Codes also specify four categories of nonmetro counties that are adjacent to metropolitan areas. These four categories were collapsed into a single 'adjacent to metro area' category. The rationale for these recodes was that in Kentucky there is limited variation in the metro and adjacent to metro categories. The remaining categories in the Urban Influence Codes were used unaltered. Previous research (Price & Reeves, in press; Reeves, 2000) had shown that this scheme for categorizing geographic location produced meaningful results when studying educational outcomes in Kentucky.[4]

Findings of the HLM Analyses

Two HLM analyses were specified to discover the short-term effects of location and student poverty on school performance and improvement. Table 2 presents the results. An elementary school in a county school district achieved an average baseline performance of 69.4. In the next bien-

[3]The discrepancy between Roeder's estimation of the annual change in accountability scores and my own can be attributed to differences of purpose and method. Roeder's purpose was descriptive rather than theoretical. His projections were designed to determine how many schools *in toto* and which schools in particular will make the 100-point proficiency target by 2014. He did not examine theorized characteristics of schools and districts that would affect the results. Furthermore, he used annual scores rather than biannual mean scores as I did. Lastly, he examined year-to-year changes in scores over a longer period of years. His analysis relied not only on CATS data but also on accountability data that was collected prior to the adoption of CATS in 1999. For my study it was important to restrict the data used to that collected by the CATS protocol.

[4]Before deciding to use this scheme, I explored the school location codes provided in the Common Core of Data (National Center for Education Statistics). My assessment of this alternative classification was that while these indicators of school location could be preferable for some purposes, they were misleading for my purpose of comparing location and poverty effects on school accountability results, at least in Kentucky. The focus of the present study was not on the precise location of the school building. Rather the location dimension was intended to indicate what sort of community the students attending a particular school lived in, as indicated by the demographic/location characteristics of the county. Because of the high proportion of nonmetro counties in Kentucky, the county may be considered the unit of local community. Furthermore, the CCD categorization scheme presented problems that militated against its use. For instance, the CCD category denoting that the school is in *a rural location inside a metropolitan area* is, for my purpose, an anomalous category. In Kentucky such a school can hardly be considered rural since it will benefit from the metropolitan tax base and its staff will benefit from higher salaries while its students will have access to urban amenities. Another handicap that discouraged me from using the CCD scheme was that some of the nonmetro categories yielded subsamples in the Kentucky data that were too small for multivariate analysis.

Table 2
Results of the HLM Analyses: Location and Student Poverty Effects on School Performance and Improvement

Fixed Effect	School performance	School improvement
Intercept	69.352 (1.925)***	4.492 (0.485)***
Independent district	3.108 (2.020)	0.109 (0.487)
District size (ln)	0.337 (0.961)	0.286 (0.194)
Per student revenue ($)	-0.003 (0.002)	-0.000 (0.000)
Nonmetro, adjacent to metro	-0.498 (1.816)	0.171 (0.411)
Nonmetro, town < 2,500	-3.211 (2.532)	1.560 (0.649)*
Nonmetro, town 2,500 – 9,999	-2.264 (1.713)	1.139 (0.437)*
Nonmetro, town ≥ 10,000	0.650 (1.722)	0.995 (0.504)*
% free/reduced lunch	-0.251 (0.028)***	-0.038 (0.029)~
% Black students (ln)	-0.051 (0.101)	0.016 (0.072)
School membership (ln)	-1.176 (0.515)*	-1.108 (0.330)**
Student/teacher ratio	0.176 (0.059)**	-0.072 (0.039)~
Middle school	-3.971 (0.910)***	-2.560 (0.458)***
High school	-8.069 (0.993)***	-2.991 (0.619)***
Combined school	-3.719 (0.820)***	-1.704 (0.568)***
1999-2000 Academic Index score	--	-0.266 (0.034)***
School-level variance component:		
Controls only model (not shown)	55.167	16.093
Full model	31.978	15.009
Net variance explained	42.0 %	6.7 %

$N_1 = 1{,}115$, $N_2 = 171$
Metric coefficient (standard error)
$\sim p < 0.10$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$

nium, the average score of the elementary school increased by 4.5 points. The performance and improvement results for the other types of schools are markedly less. For example, the high school baseline was 8.1 points below the elementary baseline, and the middle school baseline was 4.0 points less. Similarly, high schools registered an average biennial gain of only 1.5 while middle schools gained 1.9.

The baseline performance was positively influenced by the student/teacher ratio and by school membership. School membership also negatively influenced improvement, while the effect of the student/teacher ratio was negative, but only marginally significant. Total per student revenue was not significantly associated with either school performance or improvement. The variable that denoted if the district was independent proved to be nonsignificant. Nor did the percentage of Black students in the school influence either performance or improvement. On first consideration this finding may evoke surprise. A standard OLS regression analysis would show this variable to have a moderately negative effect. But standard OLS regression would not take into account the clustering of predominantly minority schools in a single, very large metropolitan school district. HLM has corrected for this clustering bias and has revealed the more generalized finding that Black students did not significantly affect either school performance or improvement when the regression estimates were aggregated district by district.

Location and poverty effects on the baseline school performance. Net of the control variables, location and student poverty explained 42% of the variance in the baseline school performance. The fixed effect of student poverty on school performance was decidedly negative. A 1% increase in students on free/reduced lunch was associated with an average 0.25 decline in the Accountability Index. The size of this effect is more apparent when a larger unit of measurement is chosen. For instance, a one standard deviation increase in poverty would lead to a 5.4-point decline in performance. Turning to nonmetro location, the effects of the dummy variables had large standard errors, and therefore were not statistically significant. In summary, the effect on school performance was almost entirely due to

student poverty and not to location, and the size of this effect was large.

Location and poverty effects on school improvement. Net of the control variables, location and student poverty explained nearly 7% of the variance in school improvement. Student poverty had a marginally significant, negative influence on improvement. A standard deviation increase in the percentage of students participating in free/reduced lunch program decreased the biennial gain by the average of 0.8. Unlike the previous analysis, here location was found to have a stronger effect. Nonmetro location produced significantly positive effects on improvement with three of the four dummy variables. The location effect was nonsignificant only in school districts that were adjacent to metro areas. Where there was a town with fewer than 2,500 inhabitants the biennial gain was predicted to increase 1.6 points above the average gain achieved by metro schools. Where the location included a town of 2,500 to 9,999 inhabitants, the increase was predicted to be 1.1 points above; and in a location with a town of 10,000 or more inhabitants the predicted increase was slightly less, at 1.0.

How equitable is Kentucky's accountability system? The results of the HLM analyses support the conclusion that location and poverty have separate effects on the short-term accountability results. Student poverty had the greater influence on school performance; location had the stronger influence on school improvement. These short-term effects are significant and raise the question of the fairness of Kentucky's accountability system. Application of the current accountability standards will, if the above results reliably indicate future trends, disproportionately underrepresent the performance of schools with higher poverty students and overrepresent the gains made by metropolitan schools.

Projecting Long-Term Effects

What are the effects of these inequities when projected over a longer period of time? This question can be answered by entering selected values into the HLM equations. Then by making certain assumptions it can be determined if schools with hypothetical characteristics are likely to reach the goal of 100, or proficiency, on the Accountability Index by the year 2014. This analysis reveals where long-term inequities may arise in Kentucky's standards-based reform, assuming that present trends continue.

Table 3 below presents the results of using the HLM equations to project the effects of poverty and location on elementary schools reaching the proficiency goal of 100 points by 2014. Since these projections were derived from the preceding HLM analysis, they are subject to the empirical and methodological limitations of that analysis. The projection analysis entails a long-range, linear extrapolation beyond the HLM results. Because of this, the projections could be considered extreme and unreasonable if there is not a basis for believing that the assumptions of a linear gain model will hold true. These assumptions must be clarified and their likelihood of stability must be assessed. Before contemplating the results of the projection analysis therefore, it is advisable to consider the assumptions that undergird it.

The projections were based on the trends extrapolated from two biennia of accountability data. The procedure specified a constant annual gain (or that the annual gain will have a constant average value) throughout the extrapolation. The estimated annual gain was derived from the HLM equation for school improvement (Table 2). As an example, a metropolitan elementary school, with average poverty and located in a county school district, was estimated to average a gain of 2.25 points each year until 2014. Because this estimate is based on only two biennia of data, its reliability could be suspect. However, the reliability of my estimates is buttressed by Roeder's (2001) calculations, which made use of longer time span of data. Combining Roeder's results with my own provides a decade's worth of data revealing an average annual gain of slightly more than 2 points on the Accountability Index. Roeder's estimate is fractionally larger than the one I use, but the difference is not enough to alter the conclusions of the present study. The story would not change, for example, if I used Roeder's estimate of annual gain instead of my own. Therefore, the estimate I am using appears to be a reasonable one based upon all of the Kentucky school improvement data that is currently available.

Other factors, if incorrectly estimated, might also disturb the projection. Suppose after collecting data for a longer span of years, student poverty and school location were found to have effects quite a bit different from those estimated in Table 2? How likely is this? The evidence for the strong negative effect of student poverty on school performance and, to a lesser degree, the negative effect on school improvement is well substantiated by prior research (Reeves, 2000; Roeder, 1999, 2000a, 2000b). The effects associated with different school locations are less well substantiated (however, see Reeves, 2000), but since the location effects appear to influence the projection to a much lesser degree than student poverty, errors in their estimation are probably not as grave. Failure to accurately estimate the effects of the control variables could also influence the projection, but here again these variables also appear to exert only small effects—with the exception of two control variables: The type of school and the previous Academic Index score were both estimated to have sizeable effects. However, the large effects associated with these two variables are also corroborated by previous studies (Roeder, 1999, 2000a).

Perhaps the greatest cause for concern about the reliability of the projection is not with the potential instability

Table 3
Long-term Effects of Location and Poverty on the Performance-based Results of Hypothetical Elementary Schools

Location and poverty characteristics		Baseline performance (2000)	Estimated annual gain	Projected performance (2014)
Free/reduced lunch = 73.0 % (+1 SD) in 1999				
Metro		63.9	1.81	89.3
Nonmetro, adjacent to metro		63.5	1.90	90.0
Nonmetro, town < 2,500		60.7	2.59	97.0
Nonmetro, town 2,500 - 9,999		61.7	2.38	95.0
Nonmetro, town ≥ 10,000		64.6	2.31	96.9
	Weighted mean	63.0	2.10	92.4
Free/reduced lunch = 51.4 % (mean) in 1999				
Metro		69.4	2.25	100.8
Nonmetro, adjacent to metro		68.9	2.33	101.5
Nonmetro, town < 2,500		66.1	3.03	108.5
Nonmetro, town 2,500 - 9,999		67.1	2.82	106.5
Nonmetro, town ≥ 10,000		70.0	2.74	108.4
	Weighted mean	68.4	2.53	103.8
Free/reduced lunch = 29.7 % (-1 SD) in 1999				
Metro		74.8	2.68	112.2
Nonmetro, adjacent to metro		74.3	2.76	112.9
Nonmetro, town < 2,500		71.5	3.46	119.9
Nonmetro, town 2,500 - 9,999		72.5	3.25	118.0
Nonmetro, town ≥ 10,000		75.4	3.17	119.8
	Weighted mean	73.8	2.96	115.3

Note. The effects shown are for hypothetical schools that are located in county districts. The values were estimated using the HLM equations and coefficients (see Appendix and Table 2).

of the variables and the effects in the model but rather with what is not in the model—the policy environment. Does educational policy in Kentucky show a tendency to change rapidly, or has it remained fairly stable? How frequently do large policy changes, of the sort that could impact the results of the projection analysis, occur? Our best guide to answering these questions is to look at what has happened in Kentucky over the last 12 years of educational reform. What we find is a concerted effort by the Kentucky Department of Education, the Kentucky Board of Education, the Kentucky General Assembly, several governors, both of the state's leading newspapers, and influential citizen groups, like the Prichard Committee for Academic Excellence, to stay the course on school accountability and to discourage making large-scale changes. Even now when Kentucky is required to align itself with the NCLB Act, there is a strong sentiment in state policy circles to limit change as much as possible in order to safeguard the integrity of Kentucky's accountability system.

The projection analysis presents only one version of an uncertain future. One can imagine if the budget crisis in Kentucky worsened, the new governor coming into office in 2004 might feel strongly tempted to make deep cuts in the state's most costly program, public education. This scenario could lead to the kind of large policy changes that would invalidate the projection, but change of this magnitude would be very difficult to justify to taxpayers and legislators in view of how much Kentucky has already invested in the current accountability system.

In devising the projections I used all coefficients at the values posted in Table 2, regardless of the level of statistical significance. The rationale for doing this was that the coefficient represented the central tendency of the effect even when it was not significant. Nonsignificance

meant, of course, that the spread of data points about the central tendency was too great to conclude with confidence that after repeated sampling the coefficient would have been greater than zero. Since the size of a nonsignificant coefficient was typically small, the effect on the projection was also small.

It will be noted that hypothetical schools have been created using the characteristics of location and student poverty (percent free/reduced lunch students). The performance of each school was estimated using the intercept that estimates the baseline performance and then adding the coefficients of the hypothesized characteristics. A similar procedure was used to calculate improvement, although the coefficient of the intercept was then halved to give the estimated annual gain. (Improvement across two biennia was a 2-year gain.) To find the projected performance in 2014, the baseline performance score was added to the product of the estimated annual gain multiplied by 14 years. For example, to obtain the projected performance of a hypothetical metro elementary school with average student poverty the following calculation was made: 69.4 + (2.25 x 14) = 100.8 (with a slight rounding error). The weighted means in Table 3 portray the effects of poverty independent of location. The means have been weighted for the proportion of schools in each type of location.

Let us now turn to the findings of the projection analysis. Looking at the projected performance results in the right-hand column of Table 3, it is apparent that the hypothetical metro elementary schools will not lead the race to the finish, despite their positions near the lead in 2000. Metro elementary schools are predicted to have fallen to the rear of the pack by 2014, regardless of student poverty level. Elementary schools in the nonmetro districts will have moved ahead of their metro counterparts. There is a projected 7.7-point difference between the school in the highest scoring location (nonmetro, town < 2,500) and the comparable metro school. The effect of student poverty is even more dramatic. The difference between the hypothetical school with the highest percentage of students in poverty and the school with the lowest percentage is 10.8 points in 2000. By 2014 the difference has grown to 22.9 points. The hypothetical elementary school where the free/reduced lunch percentage is one standard deviation above the mean is projected to fall short of the proficiency target of 100 points, regardless of location. Schools with an average percentage of students in poverty will just reach the target, and schools with the fewest students in poverty will comfortably surpass it.

These projected effects suggest the long-term influence of location and student poverty on elementary schools. A similar analysis projecting outcomes for middle schools and high schools would reveal a substantially lower success rate for these types of schools. Owing to differences between Roeder's (2001) study and my own, our findings are not directly comparable. However, the long-term projections of both studies produced results that appear to corroborate each other. For example, Roeder found that only 40% of Kentucky public schools are likely to have reached proficiency in 2014. Among the successful schools, he predicted the overwhelming majority will be elementary schools. Taking into consideration that my study employed a slightly more conservative estimate of annual gain than Roeder's, our respective findings converge remarkably well.

Are There Remedies for Inequity?

The potential for short-term and long-term inequity resulting from Kentucky's program of school assessment and accountability has been demonstrated. Unless steps are taken to change the trends that were predicted above, schools disadvantaged now will likely remain disadvantaged long into the future. What can be done to reduce these unequal educational outcomes? I will address this question by exploring a variety of remedies that are currently available, as provided by Kentucky educational policy and by the NCLB Act.

Remedies provided by KERA. High-stakes incentives are the cornerstone of Kentucky's standards-based reform. After the Kentucky Education Reform Act of 1990 was passed, the state began using financial incentives to teachers and other school personnel to encourage steady improvements in the accountability measures. A decade after these incentives were put in place they are not having the effect of transforming disadvantaged schools into super achievers (Roeder, 1999, 2000a, 2000b). Indeed, they are not designed for this purpose, because they reward performance equally, regardless of the advantages or disadvantages that different schools possess. At best, the program of rewarding improving schools allows inequality to persist. The fact that a small percentage of the schools with impoverished students have been able to achieve impressive gains over a short period of years does not erase the fact that the large majority of disadvantaged schools are not gaining nearly fast enough to shrink the distance between themselves and the schools with relatively few poor students. Schools with impoverished students that are located in rural school districts are apparently benefiting from a location effect that partially offsets the pernicious effect of student poverty.

Other remedies not yet tried. One strategy has been discussed in the state, but opposition to it has always been strong. It would entail adjusting the accountability scores by controlling for factors that disproportionately benefit some schools and disadvantage others. For example, the change in accountability scores could be statistically adjusted by controlling for student poverty and location. Advocates of this approach say that it would level the playing field and make accountability testing fair. Disadvantaged schools would no longer be penalized for their disadvan-

tage, and advantaged schools would not reap unearned bonuses. Neither KDE nor the Kentucky legislature has been willing to back such a plan, fearing that it could encourage schools to slack off in their efforts to improve.

Another remedy lies in the assistance package made available for addressing problems in schools that fail to adequately progress. KDE is authorized to assist struggling schools with school improvement funding and a scholastic review/audit. These measures may allow the lowest-improving schools to perform more successfully, but they are not contemplated to be necessary to enable the large majority of disadvantaged schools that are improving but at a slower rate than their advantaged counterparts. Providing extraordinary assistance to all disadvantaged schools that could profit from it would impose a much larger financial burden on the state (cf. Linn et al., 2002). Yet, it is precisely this remedy that may be needed if policymakers, the legislature, and taxpayers truly want to address the inequities that currently exist and could well increase as time goes by.

Other strategies to boost the performance of disadvantaged schools are probably even more unpalatable. One such strategy would be to reduce the student/teacher ratio in low-performing schools, especially in the early grades (Ehrenberg, Brewer, Gamoran, & White, 2001). But this approach, which has been implemented in several states, may be prohibitively expensive, even if limited to disadvantaged schools that are struggling. Kentucky may not have the luxury to experiment with this solution as wealthier states such as California and Tennessee have done. The U.S. Supreme Court has recently authorized another potential strategy—the use of public funds for school vouchers. Even if voucher plans turn out to be proven widely successful, this remedy will not be adopted in Kentucky any time soon because of state constitutional restrictions on the use of public funds (Blackford, 2002).

An NCLB remedy? Under the federal NCLB law, students can transfer out of schools that fail to meet their state's accountability standards for two consecutive years. The penalty applies only to schools that receive federal dollars, since these schools serve a large number of low-income students. It is a provision of the NCLB Act that could conceivably benefit disadvantaged students. But it is important to realize that this remedy is aimed at individual students in failing schools who want to improve their academic circumstances by transferring to another school that has a stronger academic record. It is difficult to see how this remedy will help the school that is failing. By allowing academically motivated students to transfer, the performance of the school will probably decline further. Moreover, it is not clear that this remedy will even benefit the students who choose to transfer. According to a recent *Lexington Herald-Leader* article (Deffendall, 2002) the choices for students who want to switch schools will be few. Transfers from a failing school will only be permitted to enroll in a school with Title I funding, and it is very likely that this school will also have low scores on the state's tests. In Kentucky rural school systems, there may not be a practical alternative to attending the local middle school and high school. At the end of the day, this provision of the NCLB Act may offer students and parents hardly more than a Hobson's choice.

Conclusion

After the effects of location and poverty were disentangled, nonmetro location was found to mitigate the negative influence that poverty has on school performance and improvement. The present study has therefore disclosed a rural advantage that does not lie in small schools and small districts alone. Although the exact source of the rural advantage was not identified, it could lie in the social capital assets that are possessed by nonmetropolitan schools, communities, and families—assets that are more often absent in core metropolitan contexts (Bryk & Schneider, 2002; Coleman, 1990; Israel et al., 2001; Khattri et al., 1997). However, it is equally important to note that this rural education advantage was found to be small in magnitude. It did not offset the negative consequences of student poverty.

The results of this study have implications for school assessment and accountability both for the state of Kentucky and for the nation. We saw that school- and district-level effects influence school performance and improvement in the short term and very possibly in the long term. Restricting assessment criteria to school improvement (i.e., gains in scores), as Kentucky does, does not guarantee that these effects are negligible and does not eliminate the potential for inequity in accountability outcomes. The evidence of the present study supports three distinct policy solutions: (a) modify testing standards and procedures so that more schools will succeed, (b) adjust school accountability scores by poverty level (and by any other relevant factors) so that disadvantaged schools are not harmed by hidden bias, or (c) be prepared to support large numbers of disadvantaged schools with a vigorous program of enrichment and assistance.

The first option handles the problem pragmatically and has already been embraced by some states (Linn et al., 2002), but it does not further the goal of excellence, even less the goal of equitable excellence. In Kentucky, state policymakers have invested a great deal in designing an accountability system that has been found meaningful and that has gained the acceptance of educators and citizens. It is not likely that policymakers will soften testing standards unless compelled to do so.

The second option may prevent disadvantaged schools from being unfairly stigmatized and sanctioned, but it will also mask the problem of inequity and make its solution

more intractable. Adjusting scores will also not receive the support of policymakers, who claim that it conveys the erroneous impression that low performing schools will be tolerated rather than compelled to change their practices. Accepting low performance would flout a central tenet of Kentucky's standards-based reform, that all students are capable of learning at high levels. Quite simply policymakers are opposed to adopting any strategy that would let low performing schools off the hook. The whole point of defining accountability standards is to create a yardstick to measure school performance and to compel low-performing schools to alter their modes of teaching so as to bring about positive change.

The third option is the path to "equitable excellence." Instead of adjusting test scores, it calls for addressing equity by maintaining realistically high standards for all schools while targeting the deficiencies of disadvantaged schools with special assistance. A variety of initiatives would be potentially useful in this effort. They include: enriching the academic culture and redesigning the curriculum of disadvantaged schools, reducing class and school size, hiring more qualified teachers of targeted subjects such as mathematics, upgrading the skills of all teachers by means of professional development, and hiring school administrators with energy and vision (Darling-Hammond & Post, 2000; Ehrenberg et al., 2001;). Standards-based reform has promoted the use of these strategies to varying degrees. But the idea has not taken hold that large numbers of disadvantaged schools may require disproportionately large infusions of these and other enhancements in order to draw near the performance levels achieved by advantaged schools.

The third option is undoubtedly the right path. But will it be prohibitively expensive for states like Kentucky to follow this path unless the federal government assumes much of the increased cost? The challenge of meeting the needs of disadvantaged students and schools is so daunting that it may require a new educational reform initiative, equivalent in its scale and impact to standards-based reform. And, in the end, closing the achievement gap between advantaged and disadvantaged schools may draw the federal government into increasing its share of school expenditures both in Kentucky and nationwide, further eroding the historic American position of reserving the lion's share of responsibility for public school finance to state and local governments.

References

American Educational Research Association. (2000). *AERA position statement concerning high-stakes testing in preK-12 education.* Retrieved July 25, 2002 from http://www.aera.net/about/policy/stakes.htm

Bickel, R., Howley, C., Williams, T., & Glascock, C. (2001). High school size, achievement equity, and cost: Robust interaction effects and tentative results. *Education Policy Analysis Archives, 9*(40). Retrieved October 9, 2001 from http://epaa.asu.edu/epaa/v9n40.html

Blackford, L. B. (2002, July 17). Vouchers unlikely soon in Kentucky. *Lexington Herald-Leader*, pp. A1, C11.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement.* New York: Russell Sage Foundation.

Catterall, J. S., Mehrens, W. A., Ryan, J. M., Flores, E. J., & Rubin, P. M. (1998). *Kentucky instructional results information system: A technical review.* Frankfort, KY: Commonwealth of Kentucky Legislative Research Commission.

Christie, K. (2002). States address achievement gaps. *Phi Delta Kappan, 84*(2), 102-103.

Coleman, J. S. (1990). Schools and the communities they serve. Equality and achievement in education (pp. 315-324). Boulder, CO: Westview Press.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Department of Education.

Darling-Hammond, L., & Post, L. (2000). Inequality in teaching and schooling: Supporting high-quality teaching and leadership in low-income schools. In R. D. Kahlenberg (Ed.), *A notion at risk: Preserving public education as an engine for social mobility* (pp. 127-167). New York: The Century Foundation Press.

Deffendall, L. (2002, July 13). Choices few for students who switch. *Lexington Herald-Leader*, pp. C1, C2.

Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & White, J. D. (2001). Does class size matter? *Scientific American, 285*(5), 78-85.

Elmore, R. F., Abelmann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington, DC: The Brookings Institution.

Fan, X., & Chen, M. J. (1999). Academic achievement of rural school students: A multi-year comparison with their peers in suburban and urban schools. *Journal of Research in Rural Education, 15*, 31-46.

Friedkin, N. E., & Necochea, J. (1988). School system size and performance: A contingency perspective. *Educational Evaluation and Policy Analysis, 10*, 237-249.

Friedman, S., & Lichter, D. T. (1998). Spatial inequality and poverty among American children. *Population Research and Policy Review, 17*, 91-109.

Gamoran, A. (2000). High standards: A strategy for equalizing opportunities to learn? In R. D. Kahlenberg (Ed.), *A notion at risk: Preserving public education as an*

engine for social mobility (pp. 93-126). New York: The Century Foundation Press.

Greenberg, E. J., & Teixeira, R. A. (1995). Nonmetro student achievement on par with metro. *Rural Development Perspectives, 10*, 17-23.

Guskey, T. R. (Ed.). (1994). *High stakes performance assessment: Perspectives on Kentucky's educational reform.* Thousand Oaks, CA: Corwin Press.

Howley, C. (1996). Compounding disadvantage: Consolidation and the effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education, 12*, 25-32.

Huang, G., & Howley, C. (1993). Mitigating disadvantage: Effects of small-scale schooling on students' achievement in Alaska. *Journal of Research in Rural Education, 9,* 137-149.

Israel, G. D., Beaulieu, L. J., & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology, 66*, 43-68.

Kahlenberg, R. D. (2000). Introduction: Making K-12 public education an engine for social mobility. In R. D. Kahlenberg (Ed.), *A notion at risk: Preserving public education as an engine for social mobility* (pp. 1-8). New York: The Century Foundation Press.

Khattri, N., Riley, K. W., & Kane, M. B. (1997). Students at risk in poor, rural areas: A review of the research. *Journal of Research in Rural Education, 13*, 79-100.

Ladd, H. F. (1996). Introduction. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: The Brookings Institution.

Lewis-Beck, M. S. (1980). *Applied regression: An introduction* (Quantitative Applications in the Social Sciences No. 22.). Thousand Oaks, CA: Sage.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3-16.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment* (J. Pellegrino, N. Chudowsky, & R. Glaser, Eds.). Washington, DC: National Academy Press.

No Child Left Behind Act of 2001, Pub. Law No. 107-110, 115 Stat. 1425 (2002). Available online at http://www.ed.gov/legislation/ESEA02/

Pelavin Research Institute. (1996). *Proceedings of the rural education issues meeting*. Washington, DC: Author.

Price, D., & Reeves, E. B. (in press). Student poverty, school accountability, and postsecondary enrollment: A challenge for educational reform in Kentucky. *Journal of Poverty.*

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd edition). Thousand Oaks, CA: Sage.

Reeves, E. B. (2000). High stakes accountability and contextual effects: An empirical study of the fairness issue. *Research in the Schools, 7*, 49-58.

Roeder, P. W. (1999). *Education reform and equitable excellence: The Kentucky experience.* Lexington, KY: RDS Publishing.

Roeder, P. W. (2000a). *Education reform and equitable excellence: The Kentucky experiment* (Update, February, 2000). Unpublished manuscript.

Roeder, P. W. (2000b). Education reform and equitable excellence: The Kentucky experiment. *Foresight, 7*(2). Retrieved July 26, 2000 from http://www.kltpre.net/Foresight/Chpt_37.htm

Roeder, P. W. (2001). *The KERA endgame: Which Kentucky schools will achieve proficiency by 2014?* Retrieved October 15, 2002 from http://www.uky.edu/~proeder/keraweb.htm

Roscigno, V. J., & Crowley, M. L. (2001). Rurality, institutional disadvantage, and achievement/attainment. *Rural Sociology, 66*, 268-292.

Stallman, J. I., & Johnson, T. G. (1996). Community factors in secondary educational achievement in Appalachia. *Youth & Society, 27*, 469-484.

Weston, S. P. (2000). *1999 Kentucky elementary school performance and poverty, 1999 Kentucky middle school performance and poverty, 1999 Kentucky high school performance and poverty: Results from the 1999 Kentucky Core Content Test.* Report prepared for The Prichard Committee for Academic Excellence Annual Meeting, July 9-10, 2000.

Young, D. J. (1998). Rural and urban differences in student achievement in science and mathematics: A multilevel analysis. *School Effectiveness and School Improvement, 9*(4), 386-418.

Appendix

Summary of the School Performance Model in Equation Format

School level model

$$Y = B0 + B1 * (\%\ \text{free/reduced lunch}) + B2 * (\text{Middle school}) + B3 * (\text{High school}) + B4 * (\text{Combined school}) + B5 * (\%\ \text{Black students[ln]}) + B6 * (\text{School student/teacher ratio}) + B7 * (\text{School membership[ln]}) + R$$

District level model

$$B0 = G00 + G01 * (\text{Independent district}) + G02 * (\text{Nonmetro, adjacent to metro}) + G03 * (\text{Nonmetro, town} < 2{,}500) + G04 * (\text{Nonmetro, town } 2{,}500 - 9{,}999) + G05 * (\text{Nonmetro, town} \geq 10{,}000) + G06 * (\text{District size[ln]}) + G07 * (\text{Per student revenue}) + U0$$

$$B1 = G10 + U1$$
$$B2 = G20$$
$$B3 = G30$$
$$B4 = G40$$
$$B5 = G50$$
$$B6 = G60$$
$$B7 = G70$$

Summary of the School Improvement Model in Equation Format

School level model

$$Y = B0 + B1 * (\%\ \text{free/reduced lunch}) + B2 * (\text{Middle school}) + B3 * (\text{High school}) + B4 * (\text{Combined school}) + B5 * (\%\ \text{Black students[ln]}) + B6 * (\text{School student/teacher ratio}) + B7 * (\text{School membership[ln]}) + B8 * (\text{1999-2000 Accountability Index score}) + R$$

District level model

$$B0 = G00 + G01 * (\text{Independent district}) + G02 * (\text{Nonmetro, adjacent to metro}) + G03 * (\text{Nonmetro, town} < 2{,}500) + G04 * (\text{Nonmetro, town } 2{,}500 - 9{,}999) + G05 * (\text{Nonmetro, town} \geq 10{,}000) + G06 * (\text{District size[ln]}) + G07 * (\text{Per student revenue}) + U0$$

$$B1 = G10 + U1$$
$$B2 = G20$$
$$B3 = G30$$
$$B4 = G40$$
$$B5 = G50$$
$$B6 = G60$$
$$B7 = G70$$
$$B8 = G80 + U8$$