

Education Policy Analysis Archives

Volume 8 Number 22

May 10, 2000

ISSN 1068-2341

A peer-reviewed scholarly electronic journal
Editor: Gene V Glass, College of Education
Arizona State University

Copyright 2000, the **EDUCATION POLICY ANALYSIS ARCHIVES**.
Permission is hereby granted to copy any article
if **EPAA** is credited and copies are not sold.

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the [ERIC Clearinghouse on Assessment and Evaluation](#) and are permanently archived in *Resources in Education*.

The Influence of Scale on School Performance: A Multi-Level Extension of the Matthew Principle

Robert Bickel
Marshall University

Craig Howley
Ohio University and AEL, Inc.

Abstract

In this study, we investigate the joint influence of school and district size on school performance among schools with eighth grades (n=367) and schools with eleventh grades in Georgia (n=298). Schools are the unit of analysis in this study because schools are increasingly the unit on which states fix the responsibility to be accountable. The methodology further develops investigations along the line of evidence suggesting that the influence of size is contingent on socioeconomic status (SES). All previous studies have used a single-level regression model (i.e., schools or districts). This study confronts the issue of cross-level interaction of SES and size (i.e., schools and districts) with a single-equation-relative-effects model to interpret the joint influence of school and district size on school performance (i.e., the dependent variable is a school-level variable). It also tests the equity of school-level

outcomes jointly by school and district size. Georgia was chosen for study because previous single-level analysis there had revealed no influence of district size on performance (measured at the district level). Findings from this study show substantial cross-level influences of school and district size at the 8th grade, and weaker influences at the 11th grade. The equity effects, however, are strong at both grade levels and show a distinctive pattern of size interactions. Results are interpreted to draw implications for a "structuralist" view of school and district restructuring, with particular concern for schooling to serve impoverished communities. The authors argue the importance of a notion of "scaling" in the system of schooling, advocating the particular need to create smaller districts as well as smaller schools as a route to both school excellence and equity of school outcomes.

Research on the role of school and district size as an influence on school performance has a long history and a large literature (see, for example, Barker & Gump, 1964; Guthrie, 1979; McDill, Natriello, & Pallas, 1986; Smith & DeYoung, 1988; Fowler, 1991; Walberg & Walberg, 1994; Khattari, Riley & Kane, 1997; Stiefel, Berne, Iatarola, & Fruchter, 2000). The varying methods used to study the issue have, of course, generated conflicting results (Rossmiller, 1987; Caldas, 1993; Lamdin, 1995; Rivkin, Hanushek & Kain, 1998). In consequence, size has often been relegated to the status of an obligatory but uninteresting control variable. Not infrequently, it has simply been ignored altogether (Barr & Dreeben, 1983; Burtless, 1996; Gamoran & Dreeben, 1986; Farkas, 1996; Wyatt, 1996; Hanushek, 1997, 1998). A recent school effectiveness review by eleven production-function virtuosos, for example, devoted just three of its 396 pages to school size (Betts, 1996, pp. 166-168). Consequences of variability in school size, moreover, were, in passing, judged to be uncertain. District size is considered even less interesting than school size by most researchers interested in school performance.

The study reported here, by contrast, builds on a line of evidence that has related the size of both districts and schools to aggregate student achievement. Previous research developing this line of evidence, however, has constructed only single-level analyses (schools or districts). The present study deploys a multi-level method (Boyd & Iversen, 1979; Iversen, 1991) to link effects at the two levels. In other words, this new work constitutes a first step from an empirical consideration of "size effects" toward an empirical consideration of "scale effects" (cf. Guthrie, 1979).

School System Scale: A Timely Issue

A great deal of skepticism exists about the role of size as a structural condition of US schooling. Educators have generally disparaged the role of structure and focused attention on the role of process. This focus of interest is easy to fathom. Both school teachers and administrators devote themselves to the processes of teaching and administration; the structural features of their practices are, for the most part, tacit. Teachers and principals encounter schools and districts as the particular stages on which they personally enact their work and deploy professional processes. Whatever structural variety might distinguish one such "stage" from the next, teachers and principals do not often personally experience it. Superintendents, by contrast, are better positioned to develop a sense of structural differences among schools and districts, but such an

appreciation might be almost as exceptional among superintendents as it is among other educators, since process also consumes most of a superintendent's time.

This propensity to focus on process has a philosophical dimension, as well. A structuralist view confines free will to an apparently smaller range of influence as compared to a view that privileges process. Education, and the culture of education, pays considerable homage to free will (cf. Bruner, 1996). In the grandest tradition, education is seen as the route to a "larger life" open to everyone equally (e.g., Prichard Committee, 1990). James Coleman was among the first to point out that equal educational opportunity was more problematic than previously imagined, of course, and due to structural reasons. The school effectiveness literature ensued and dramatically valorized process as the profession's response to a sociological perspective on structure; school reform has had a procedural focus ever since (cf. Dorn, 1998).

Recent research and current events, however, have combined to challenge the conventional disposition to privilege process over structure. First, nearly a decade of research on school size (in particular) has developed a preponderance of evidence to suggest that smaller school size would improve schooling in impoverished communities (Howley, 1989; Irmsher, 1997; Raywid, 1999). Second, school-shooting tragedies have curiously and sadly brought the issue of school size to popular attention. Possibly as a result of these awful events, the US Secretary of Education and the Governors of Georgia and North Carolina have recently spoken in favor of small schools. Surprisingly, the Secretary praised the resistance of rural communities that have fought fiercely for decades to preserve their small schools in the face of consolidation (Riley, 1999). It has, of course, been a losing battle, with some fortunate exceptions.

The recent attention has not even begun to challenge the privileged position that process enjoys, of course, and many observers continue to believe that administrative arrangements like "schools-within-schools" and "houses" can replicate the processes presumed to characterize small scale. Both Mary Anne Raywid (1996) and Deborah Meier (1995) argue persuasively that the conditions of smallness entail characteristics tantamount to structural difference: separate administration, separate budgets, distinctive authority, unique cultures, and so forth. Simulations, it turns out, have difficulty reproducing these structural features of small scale.

Nonetheless the rhetorical change is itself dramatic. No longer does size appear merely as a footnote to effectiveness studies or as a container of essentially interesting processes, but as a distinct phenomenon. School size now matters in discourse, anyhow.

School district size, however, continues to be regarded as a much less interesting issue than school size. The size of a district would seem to have no direct and little if any net influence on student achievement. As a variable, district size seems quite remote from student learning. Thus, most studies have considered district size almost purely as an administrative issue bearing on resource allocation (e.g., Bidwell & Kasarda, 1975; Meyer, Scott, & Strang, 1987). There have been a few exceptions within these studies, of course. Bidwell and Kasarda (1975) studied district size and concluded its influence on school performance was complex and contradictory:

The total effects of [district] size were slight because its consequences for output, transmitted mainly by the structural and staff qualifications variables, were of roughly equal strength in a positive and in a negative direction.... It was associated with well-qualified staff and low administrative intensity (and, therefore, we have argued, with minimal diversion of human resources away from front-line tasks). But large size also meant more students to teach and thus higher ratios of students to

teachers. (p. 69)

However, beginning with a 1988 study (Friedkin & Necochea, 1988), a new line of evidence has developed the hypothesis that the influence of both school and district size on aggregate performance is contingent on socioeconomic status. The direction of the effect has implicated small size (of schools and districts separately analyzed) as productive for the performance of schools or districts serving more impoverished communities, but larger size as productive for more affluent communities. Howley (1996) replicated the California study in West Virginia and reported similar results. Recent work (to be considered shortly) has extended the single-level findings to Georgia, Montana, Ohio, and Texas—with nearly identical results.

Relevant Literature

Researchers' tendency to overlook the interaction of school and district size with other variables (such as poverty) may be a disabling limitation of most studies that investigate the influence of school and district size on achievement, including quite recent efforts (e.g., Stiefel et al., 2000; Mik & Flynn, 1996; Riordan, 1997). This oversight tends to perpetuate the view that one size must fit all circumstances, or that some universally "best size" must exist (e.g., Lee & Smith, 1997; Stevenson, 1996). On this dubious view, size-related benefits and size-related costs are inadvertently construed as being enjoyed equally by all students (Conant, 1959; Haller, 1992; Haller, Monk, & Tien, 1993; Hemmings, 1996). Stiefel and colleagues (2000), using a somewhat more refreshing approach, recently found that small regular 9-12 high schools have a budget-per-graduate that is no greater than the budget-per-graduate of other 9-12 high schools, and, in some cases a much cheaper budget-per-graduate. (The Berne study, however, uses a small sample of schools from a single large city (n=121) and leaves aside the question of the difference between budgeted and actual costs. The conclusions about small school size, unfortunately, rest on data from just 19 small high schools, of which only 8 are "regular" schools!)

Within the past decade, however, a growing body of empirical research has held that size is negatively associated with most measures of educational productivity. These conclusions encompass measured achievement levels, dropout rates, grade retention rates, and college enrollment rates (e.g., Walberg & Walberg, 1994; Stevens & Peltier, 1995; Fowler, 1995; Mik & Flynn, 1996). The drift of the past decade of this research, then, is to portray the optimal or best size as somewhat smaller than it was after James Conant proposed 400 students as the absolute minimum size for a suitably "comprehensive" high school (Conant, 1959; Lee & Smith, 1997).

Seldom have policy makers or researchers asked "Better for whom?" or "Better for what?" or "Better under what conditions?" Asking such questions, of course, may be seen as leading to unbearable complications. Again, in this welter of interest, indifference, and outright evasion, the role of district size is seldom considered, though both Herbert Walberg's (urban) and John Alspaugh's work (rural) remain notable exceptions (e.g., Alspaugh, 1995; Walberg & Walberg, 1994).

Size-by-Socioeconomic Status Interaction Effects

The joint or interactive, rather than independent, effects of size and socioeconomic status (SES), may also have contributed to renewed interest in smaller schools and

districts. If smaller schools and districts are shown to benefit some settings, the new conventional wisdom (i.e., "smaller is better") gains support.

Specifically, interaction effects reported in some studies suggest that the well-known adverse consequences of poverty are tied to school size and, to some extent to district size, in substantively important ways. In brief, as size increases, the mean achievement of a school or district with less-advantaged students declines. The greater the concentration of less-advantaged students attending a school, the steeper the decline.

Investigations of the interaction hypothesis are relatively new, and multiple replications have only recently been undertaken and completed (see Howley & Bickel, 1999, for a recent synthesis of results in four states). Replications are important because without them, confidence in findings would be comparatively weak; research done in other locations could well yield different, and perhaps sharply conflicting, results.

The additional replications, however, now extend the scope of findings to Georgia (Bickel, 1999a), Montana (Howley, 1999a), Ohio (Howley, 1999b), and Texas (Bickel, 1999b). Previous work concerned California (Friedkin & Necochea, 1988); Alaska (Huang & Howley, 1993, in a study in which students were the unit of analysis), and West Virginia (Howley, 1996). These states represent considerable variety salient to the structure and operation of schooling in the United States—rural and urban mix, ethnic mix, magnitude of influence of State Education Agency, district organization types, school and district size, and funding inequity (Howley & Bickel, 1999).

The school-level findings in these single-level analyses are robust. In every study, an interaction effect has been confirmed. The effect varies from very strong (California, Georgia, Ohio, Texas, and West Virginia) to weak, (Montana) (Note 1). The overall conclusion is that smaller schools help maximize achievement for schools serving impoverished communities, but that larger schools serve the same function for more affluent communities.

Robust district-level interaction effects, however, were discovered in the four recent studies only in Ohio. Somewhat weaker direct negative effects of district size were reported for Texas; still weaker direct and interactive effects were evident in Montana. No district-level interactions were found in the Georgia study (Bickel, 1999a). The recent findings about district-level effects differed from the earlier findings for California and West Virginia, where substantial district-level interactions were evident (Friedkin & Necochea, 1988; Howley, 1996).

Equity Effects

In addition to reviving interest in school size as a variable of importance in educational research, this work has begun to sensitize researchers, policymakers, journalists, and (perhaps most notably) citizens to equity concerns associated with school size. One-size-fits-all is no longer a unanimous judgment. Some researchers and policymakers have indeed begun to ask, "Best-size-for-whom?" (Henderson & Raywid, 1994; Devine, 1996).

In the five replications of the Friedkin and Necochea work (i.e., West Virginia, Georgia, Montana, Ohio, and Texas) Howley and Bickel also hypothesized equity effects of size. This hypothesis proceeds logically from confirmation of the interaction hypothesis. Namely, if small size improves the odds of academic success in small schools and districts (a sort of "excellence effect" of size), then the usual relationship between SES and performance must be to some extent disrupted in them as compared to larger schools and districts. Simple zero-order correlational analysis was used to measure the magnitude of relationship between SES and achievement in smaller versus

larger units (schools or districts divided at the median in these separate data sets).

The equity effects of size are more consistent and more impressive, in fact, than the excellence effects. At all grade levels, in all five states, for both schools and districts, for a variety of alternative measures of SES, and for quite different sorts of achievement tests (i.e., both criterion-referenced and norm-referenced), the amount of variance in achievement associated with SES is substantially reduced in smaller units. In most cases, the magnitude of the relationship (Note 2) among the smaller units is about half what it is among the larger units (Howley, 1996; Howley & Bickel, 1999).

The Challenge of Cross-Level Interactions

Although the "excellence effects" of school size and the "equity effects" of both school and district size seem clear from the analyses reported by Howley and Bickel (1999), failure to confirm interaction "excellence effects" for districts in some states is intriguing. The line of evidence about school and district size has not, however, thus far included examinations of possible links between school size and district size. As a result, if unacknowledged multi-level contextual effects were present, previous studies would have ignored some portion of the structural influence of size on achievement. If the cultivation of high levels of achievement is a complex matter dependent on multiple influences, then we ought to suspect the existence of cross-level influences.

Further, discovery of such cross-level influences could be considered evidence that a structural notion of organizational scale was relevant to the enterprise of schooling—most particularly to the cultivation of academic achievement. If such cross-level relationships existed, administrators and policy makers would be well advised to coordinate their view of school size with a view of district size—and eventually with classroom size, and individual student performance, at one end of the spectrum, and size of the state and even national systems at the other end. The phenomenon of scaling could be seen as a structural characteristic of state school systems (see Thiétart & Forgues, 1995, for an interesting discussion of scaling as a feature of nonlinear dynamic systems in a chaotic state).

Methods

The present study addresses these issues by extending the consideration of "excellence effects" and "equity effects" of school and district size to a multi-level analysis with cross-level interaction terms. We chose to examine these relationships with the data for Georgia precisely because no effects of district size—either direct or interactive—had been discovered in the single-level analyses conducted by Bickel (1999a). On the basis of district-level effects that are inconsistently evident across states, we hypothesize the presence of cross-level interactions that could not be detected in the previous single-level analysis.

The Georgia dataset on which all analyses in this report are based is available for download here in any one of three formats:

- [SPSS \(409K filesize\)](#),
- [Excell \(1.65M\)](#), or
- [ASCII text \(460K\)](#).

We might as easily have chosen any of the other states, but the use of individual states is advisable for two reasons, the first theoretical and the second practical. First, from the perspective of scale, each state constitutes a uniquely structured system. In this sense, combining dissimilar states is more likely to misrepresent reality than to provide a fuller picture of it. Second, since comparable achievement measures are not available for schools and districts across the four states for which we have assembled recent data, the merging of data sets would necessarily inflate measurement error.

A Single-Equation Relative-Effects Model

To study further previously identified equity effects, we specifically ask, in this two-level analysis, if there are cross-level interaction effects that remain significant in regression equations constructed to include school and district size, as well as school and district SES, and which also control for the proportion of students who are African American, the proportion of students from ethnic minorities, and pupil-teacher ratio (a proxy for class size). Our focal interaction terms are the products of (1) district size and school SES and (2) school size and district SES. Our model also includes the two original interaction terms: (1) the product of district size and district SES and (2) the product of school size and school SES.

We use a procedure developed by Boyd and Iversen (1979) and Iversen (1991). It employs ordinary least squares estimates (Note 3) of partial regression coefficients for school-level variables, district-level variables, and school-by-district interactions in the same equation. In effect, we are combining school-level and district-level regression models, and including school-by-district interactions, which reflect variability in district-level effects from school to school (Bryk & Raudenbush, 1992, pp. 70-74). The dependent variables in these equations are always school-level performance measures.

We adopt the single-equation relative-effects version of the model, since school-level and district-level variables are likely to be closely correlated. In this model, school-level variables are centered with respect to their group means (i.e., district means) and district-level variables are centered with respect to the grand mean. Centering all independent variables in this way helps to avoid inflated estimates of standard errors due to multicollinearity (Cronbach, 1987). Centering also enables us to unambiguously partition the percentage of variance in a dependent variable accounted for by each set of independent variables in our multilevel models (Iversen, 1991). Four such distinct sets of independent variables exist in our model: (1) the set of individual-level (school) variables, (2) the set of group-level (district) variables, (3) the set of single-variable interactions by level (e.g., the product of school size and district size), and (4) a set of within and cross-level interactions of different variables. Within the fourth set of variables are found the focal interactions of this study—the two cross-level interactions of SES and size: (1) the product of district size and school SES and (2) the product school size and district SES.

Examination of residuals plotted against the independent variables shows that the residuals are not uniformly distributed with respect to SPANSIZE for the 8th grade outcome measures. The same is true for FREEPCT when using the eleventh grade outcome measures. As a result, we used weighted least squares to remedy these departures from homoscedasticity, thereby restoring the efficiency of the estimators (Gujurati, 1995, pp. 381-390).

Data Sources and Variables

Official representations describe Georgia as a state with an educational system encompassing approximately 1800 public schools (e.g., Georgia Department of Education, 1999). The data set we are using, for school year 1996-97, contains complete information on 1626 regular public schools. For this study we selected for analysis data about the universe of schools with grade 8 or grade 11 test scores. Grade 8 is the grade level in Georgia with scores prior to the wave of early-school leaving that transpires at the high school level (generally grade 10), whereas grade 11 data portray the relationships that prevail subsequent to this too-familiar exodus.

The choice of these grade levels for analysis is therefore strategic. First, students from impoverished backgrounds become dropouts more frequently than students from more affluent backgrounds. Second, this being the case, the demography of schooling at grade 11 will differ somewhat from the demography at grade 8, namely in the fact that the proportion of impoverished students will have declined. Third, the probable effect of these changed conditions, we hypothesize, will be to weaken grade 11 results. The reason for this inference is that if smaller sizes positively influence achievement in impoverished schools, demographic changes in larger schools serving impoverished students will, in effect, cast off the cause of their negative influence—by removing disproportionate numbers of impoverished students. (Note 4)

Dependent variables. Dependent variables are school-level percentile rank scores for eight subtests of the widely used Iowa Test of Basic Skills (grade 8) and school-level percentage of students passing the first administration of the Georgia High School Graduation Test (grade 11). School-level means vary dramatically with both tests, from as low as the first percentile to as high as 93rd for the ITBS and from 11 to 100 percent passing (on the grade 11 Graduation Test).

Seven of the ITBS subtests are designed to measure achievement in reading comprehension, mathematics, reading vocabulary, social studies, language arts, science, and research skills. The eighth subtest is a composite measure, intended to provide a global gauge of achievement.

The High School Graduation Test is used in this study because the ITBS is not administered above grade 8 in Georgia. The Graduation test gauges achievement in English, mathematics, social studies, and science. In addition, students receive a composite score. First administration passing percentages for the five scores are used as our outcome measures for the eleventh grade.

Independent variables. Our main predictor variables, (each measured at the school level, at the district level, and as the interaction between the school and district level) include the following: (1) number of students per grade level in thousand-student units as our measure of size (SPANSIZE); (2) proportion of all students eligible for free or reduced-price meals (FREEPCT); (3) proportion of African-American students (BLACKPCT); (4) proportion minority (i.e., nonwhite) students (MINORPCT); and (5) student-teacher ratio (S/SRATIO), a proxy for class size. We include student-ratio, in particular, to address the possibility that any findings might principally be the result of differences in class size, rather than differences in school or district size.

In order to test for the existence of cross-level interactions between size and SES, we include four interaction terms: (1) school SPANSIZE by school FREEPCT, which is the same as the school-level interaction term that had proven significant in previous single-level analyses; (2) district SPANSIZE by district FREEPCT, which is the same as the district-level interaction term that had proven non-significant in previous single-level analyses of Georgia data; (3) district SPANSIZE by school FREEPCT, which is one cross-level interaction term of interest in this multi-level analysis; and (4) school SPANSIZE by district FREEPCT, the other cross-level interaction term of interest in the

present study.

Results

Tables 1 and 2 provide descriptive statistics (means and standard deviations) for our dependent and independent variables for grade 8 and 11, respectively. SPANSIZE, at both the school and district level is measured in units of 1,000 students. A standard deviation of ".NNN," in the case of district size, for instance, is therefore equivalent to the product of ".NNN" and 1,000. Tables 3 through 10 report regression results (Note 5) for the eight achievement measures that predict school performance at the 8th grade level. The first panel in each table apportions explained variance in three columns to (1) individual-level (school-level), (2) group-level (district-level), and (3) individual-by-group (school by district) interactions. The second panel reports, in a single column, the variance attributable to interactions among SES and size variables, at both levels (i.e., individual and group), yielding the four interaction terms specified in the concluding paragraph of the methods section.

In the reporting of results below, only selected tables are presented, which nonetheless convey the findings from the complete set of analyses. [The complete set of tables in Rich Text Format can be downloaded from this point.](#)

Table 1

Descriptive Statistics: Grade 8

Dependent Variables Schools

	Mean	St. Dev.
READING COMPREHENSION	47.02	12.88
MATHEMATICS	52.26	12.42
READING VOCABULARY	43.82	15.05
LANGUAGE ARTS	54.20	12.72
SOCIAL STUDIES	51.31	12.04
SCIENCE	51.07	13.88
RESEARCH SKILLS	53.01	12.60
COMPOSITE	51.25	13.71

Independent Variables

	Mean/(St. Dev.)	
	Districts	Schools
SPANSIZE	0.219 (0.101)	0.259 (0.124)
FREEPCT	48.18 (17.48)	45.28 (22.93)
BLACKPCT	34.47 (25.25)	37.29 (29.66)

MINORPCT	2.91 (4.22)	4.14 (5.41)
S/RRATIO	16.13 (1.51)	16.25 (1.86)
	N=158	N=367

Table 2
Descriptive Statistics: Grade 11

Dependent Variables		
Schools		
	Mean	St. Dev.
ENGLISH	92.87	5.18
MATHEMATICS	85.33	9.77
SCIENCE	70.66	15.22
SOCIAL STUDIES	75.14	12.97
COMPOSITE	63.89	16.41
Independent Variables		
	Mean/(St. Dev.)	
	Districts	Schools
SPANSIZE	0.233 (0.139)	0.280 (0.114)
FREEPCT	48.18 (19.76)	33.49 (21.26)
BLACKPCT	35.42 (25.30)	38.03 (29.80)
MINORPCT	2.53 (3.42)	3.84 (5.03)
S/RRATIO	17.03 (2.34)	17.74 (3.15)
	N=155	N=298

Recall that previous single-level analyses reported statistically significant and negative SPANSIZE by FREEPCT interaction effects. These conspicuous effects meant that as school (and in some states, district) size increased, the mean achievement costs associated with less-advantaged students increased. Tables 1 through 8 again confirm interaction effects, but the interactions portrayed there are quite clearly shown to represent a complex phenomenon that escaped notice in single-level analyses. These more complex effects were predictably masked in the earlier single-level analyses, since those analyses examined schools and districts separately. The following written report of the findings may be difficult to follow, but the Tables themselves actually picture a

consistently complex set of relationships prevailing between schools and districts as those complex relationships influence school-level performance. We encourage readers to refer to the Tables as they read the following discussion.

Eighth Grade "Excellence Effects"

Combining schools and districts in a multilevel analysis, the single-level SPANSIZE by FREEPCT interaction effects that were so conspicuous in the previous single-level research are not evident at all at the 8th grade. However, several interesting (and uniquely specified) single-level and cross-level interactions are present in the equations. Overall this means that the effects of size on achievement depend on multiple influences, and not merely school- or district-level SES. One size is shown more clearly than ever before not to fit all cases, and, at the same time, these results suggest that the influential features of circumstance vary to such an extent that each setting can be understood as unique. We present this conclusion prematurely in order to help readers take a wider perspective on the presentation of detailed findings that follows.

Single Variables Within and Across Levels. First let us consider the results given in panel 1 of Tables 3 through 10 (the unique influence of single variables at each of two levels separately and then jointly across levels). We will interpret the results of Table 10 (composite achievement) only, as the results given there can be viewed as not only encompassing the generality of the findings reported in Tables 3 through 9, but as representing a summative indicator of school performance. Readers are, however, directed to those other Tables to observe the somewhat variant results among the various ITBS subtests. We will first consider the single variables as unique school-level and district-level influences (Note 6):

(1) Both FREEPCT (-) and BLACKPCT (-) exhibit uniquely significant ($p < .001$ and $< .01$, respectively) school-level influences in the equation, accounting for 26.4% of the variance in school-level performance. Neither SPANSIZE nor S/RRATIO (our proxy for class size) show any net direct influence at the school level.

(2) FREEPCT (-) and MINORPCT (+) exhibit uniquely significant ($p < .001$ and $p < .01$, respectively) district-level influences in the equation, accounting for 31.3% of the variance in school performance.

Table 10
Weighted Regression Results with Corrected Standard Error
Grade 8: Composite Score

Unstandardized and (Standardized) Regression Coefficients

	Individual-Level	Group-Level	Individual by Group Interactions
SPANSIZE	-6.401 (-.050)	27.174 (.094)	-308.619** (-.167)

FREEPCT	-0.401*** (-.418)	-0.340*** (-.460)	-0.005* (-.098)
BLACKPCT	-0.119** (-.207)	-0.001 (-.002)	-0.003** (-.212)
MINORPCT	0.112 (.040)	0.347** (.121)	0.014 (.033)
S/RRATIO	0.457 (.043)	-0.660 (-.060)	-0.460 (-.068)
Variance Explained	26.4%	31.3%	10.8%

Within-Level and Cross-Level Interactions

SCHOOL SPANSIZE by SCHOOL FREEPCT	0.141 (.024)
DISTRICT SPANSIZE by DISTRICT FREEPCT	-0.332 (-.023)
DISTRICT SPANSIZE by SCHOOL FREEPCT	-4.304*** (-.211)
SCHOOL SPANSIZE by DISTRICT FREEPCT	-1.046*** (-.237)
Variance Explained	10.7%

Residual Intraclass Correlation .056
School/District Ratio 2.32
Standard Error Inflation 6.88% (Corrected)

Partial Derivatives for Y with Respect to (1) SCHOOL SPANSIZE and (2) DISTRICT SPANSIZE

$$Y \text{ wrt } 1 = - 308.619 \times (\text{DISTRICT SPANSIZE}) - 1.046 \times (\text{DISTRICT FREEPCT})$$

$$Y \text{ wrt } 2 = - 308.619 \times (\text{SCHOOL SPANSIZE}) - 4.304 \times (\text{SCHOOL FREEPCT})$$

*p <.05

** p <.01

***p <.001

These two single-level results show that a substantial portion of the variance in school performance (i.e., mean ITBS percentile rank in a school) actually is accounted for by district-level influences. Poverty contributes a negative influence that is about 4 times the magnitude of the positive influence of MINORPCT. The direct influence of district size and district student teacher ratio, we note, are once again nonsignificant.

We next consider the individual by group interactions reported in column 3 of panel 1 (Table 10). This column reports cross-level interactions for each of the major variables separately. That is, these reported interactions compute the interactive (joint) influence of SPANSIZE, FREEPCT, BLACKPCT, MINORPCT, and S/SRATIO at the two levels. Results, which account for a unique 10.8% of the variance in school-level performance, are summarized as follows:

- (1) The unique interactive influence, across levels, of SPANSIZE (-) is highly significant (p<.001).

(2) The unique interactive influence, across levels, of FREEPCT (-) is somewhat significant ($p < .05$).

(3) The unique interactive influence, across levels, of BLACKPCT (-) is also significant ($p < .01$).

(4) There is no unique interactive influence, across levels, of MINORPCT or S/RRATIO.

To interpret these interactive results, recall that all independent variables are centered for the regression analyses. Values of the variables that fall below the mean are negative and values that fall above the mean are positive. The product of two negative values at the district level (e.g., low district poverty) and school level (small school size) will yield positive values of the interactive variable, just as the product of positive values at both levels will yield positive results. In this Georgia data set, the existence of small schools in small districts, and the existence of large schools in large districts are conditions uniquely associated with lower school performance. (Note 7) Similar inferences can be drawn in the case of FREEPCT (though the influence here accounts uniquely for less than 1% of school performance) and BLACKPCT. It is crucial for readers to keep in mind that the influences on school performance discussed thus far are not interpretable in isolation from the totality of size influences. This research is developing a model of cross-level influence of size on school performance. In this model, however, we can see that single-variable influences within and across levels account for almost 70% of the variance in school performance.

Variables Interacting Within and Across Levels The single variables—whether uniquely at different levels, or jointly across levels—present a substantial but still incomplete view of influences on school performance. These influences, in this analysis, are completed by an analysis of interactions between variables, both within and across levels. We turn next, therefore, to a consideration of these influences, given in the second panel of Tables 3 through 10. Again, discussion centers on Table 10 (composite achievement) which, in the case of interactions between pairs of focal variables (SES and size), very closely parallels results presented in Tables 3 through 9. We observe the following results (again, directionality is given parenthetically):

(1) The single-level interactions of FREEPCT and SPANSIZE, whether school- or district-level influences, are not statistically significant.

(2) The interaction (-) of SPANSIZE as a district-level influence and FREEPCT as a school- level influence is highly significant ($p < .001$).

(3) The interaction (-) of SPANSIZE as a school- level influence and FREEPCT as a district-level influence is highly significant ($p < .001$).

The two significant interactions together account for an additional 10.7% in the variation of school performance. Thus, the two-level model accounts for 79.2% of the variance in the performance of Georgia schools with an 8th grade. In other words, just 20% of the variance in school performance is the result of other influences—including school processes (such matters as curriculum and instruction).

The first interaction, the statistically significant and negative interaction of district-level SPANSIZE by school- level FREEPCT, shows two things. First, as district

sizes increase, the mean achievement cost associated with increases in the proportion of less-advantaged students at the school level increases as well. (Note 8) Second—as in the previously reported single-level analyses—the converse also pertains: As district sizes decrease (negative values of district size as a centered variable), the mean achievement cost associated with decreases in the proportion of less-advantaged students (i.e., negative values on school-level poverty) at the school level increases as well. In other words, more affluent school-communities appear to be better served by being in larger districts, but less affluent school-communities appear to be better served by being in smaller districts. Put most simply, district poverty and large school size are shown to jointly hurt predicted school-level performance, just as district affluence and small school size are shown to do. The relationship is interactive—it cuts two ways.

The second interaction, the statistically significant and negative interaction of school SPANSIZE by district FREEPCT follows the preceding interpretation. First, as school sizes increase, the mean achievement cost associated with increases in the proportion of less-advantaged students at the district level also increases. Second, as above, the converse is true as well: As school sizes decrease, the mean achievement cost associated with being in a district with decreases in the proportion of less-advantaged students also increases. The simple form of this statement, again, would be: school poverty and large district size are shown to hurt predicted school-level performance, just as school affluence and small district size are shown to do. Again, this interactive relationship cuts two ways

Eleventh Grade "Excellence Effects"

Tables 11-15 present the regression results using the five eleventh grade outcome measures. As predicted, the 11th grade results are less consistent than the 8th grade regressions (Tables 3 through 10). Interestingly, the cross-level interaction of school SPANSIZE by district FREEPCT is highly statistically significant, alone accounts for as much as 15% of the variance in school-level performance, and exhibits the expected negative sign in each equation. As with the 8th grade results, this means that as school sizes increase, the mean achievement cost associated with being in districts with increasingly less-advantaged students also increases. As before, large schools in low-income districts encounter a decided achievement disadvantage. Overall, the 11th grade "excellence" effects of size are considerably muted, and they leave their mark most particularly with the cross-level interaction of SPANSIZE and FREEPCT. (Note 9)

In general, the 11th grade results account for less variance than the 8th grade results. In the case of the composite score (Table 15), for instance, the model explains about 50% of the variance in school-level performance. The greatest proportion of variance accounted for by our model appears for mathematics (about 66%); the low is English (less than 30%). Mathematics, we observe, is a highly differentiated school subject at the high-school level, with the first course in algebra serving in the famous "gatekeeper" role (Silva & Moses, 1990) (Note 10). In other words, structural influences (poverty, race, size and the interactions among them) might exert a stronger influence on school performance than they would in less differentiated subjects such as English.

Table 15

Weighted Regression Results with Corrected Standard Error

Grade 11: Composite Score

Unstandardized and (Standardized) Regression Coefficients

	Individual-Level	Group-Level	Individual by Group Interactions
SPANSIZE	3.688 (.027)	19.952 (.052)	-133.985 (-.100)
FREEPCT	-0.413*** (-.304)	-0.187* (-.222)	-0.003 (-.066)
BLACKPCT	-0.257*** (-.378)	-0.116** (-.206)	-0.001 (-.067)
MINORPCT	0.321 (.088)	0.262 (.070)	-0.015 (-.028)
S/RRATIO	-0.915* (-.126)	-0.145 (-.013)	0.096 (.060)
Variance Explained	28.7%	10.0%	1.7%

Within-Level and Cross-Level Interactions

SCHOOL SPANSIZE by SCHOOL FREEPCT	0.281 (-.075)
DISTRICT SPANSIZE by DISTRICT FREEPCT	-0.423 (-.025)
DISTRICT SPANSIZE by SCHOOL FREEPCT	-0.027 (-.001)
SCHOOL SPANSIZE by DISTRICT FREEPCT	-1.357*** (-.456)
Variance Explained	8.0%

Residual Intraclass Correlation .066
 School/District Ratio 1.96
 Standard Error Inflation 5.96% (Corrected)

Partial Derivatives for Y with Respect to (1) SCHOOL SPANSIZE and (2) DISTRICT SPANSIZE

Y wrt 1= - 1.357 x (DISTRICT SPANSIZE)
 Y wrt 2 = COEFFICIENTS NOT STATISTICALLY SIGNIFICANT

*p <.05
 ** p <.01
 ***p <.001

Interpreting the Effect Sizes of Size

The regression equations provide a prospective tool with which to estimate the effects of projected changes in size (of schools and districts) on school performance in

Georgia relevant to the independent variables that describe a school's context. In order to interpret these predicted effects of size on school performance, we adapt the technique pioneered by Friedkin and Necochea (1988).

Those researchers differentiated their regression equations in order to infer a rate of change in achievement attributable to size, relative to a school's or district's poverty level. Their procedure found the partial derivative (Note 11) of school or district performance with respect to socioeconomic status. The partial derivative was then evaluated to find the rate of achievement change associated with changes in school or district size for schools or districts of a certain SES. This is the technique also used in the work recently reported by Bickel and Howley (e.g., Howley & Bickel, 1999).

Since our goal here is to provide a fuller quantitative account of the relationship between size and SES we have computed partial derivatives of the regression equations that give the rate of change in the dependent variable (school performance) with respect to size (school or district), holding poverty (FREEPCT) constant (at two levels of influence). It is important to remember that the dependent variable in the partial derivatives represents a rate: change in school performance per change in size.

Because this is a two-level analysis, however, two equations are necessary. One equation describes the predicted influence of changes in school size on school performance, and in this analysis that rate turns out to be a function of district-level variables. The other equation describes the predicted influence of changes in district size on school performance (in this case as a function of school-level variables). Think of this relationship as follows: Y wrt 1 is a rate of change in school performance per change in the size of a school. But this rate, in cross-level analysis, depends on district-level characteristics. Y wrt 2 is a rate of change in school performance per change in district size; this rate depends in cross-level analysis on school-level characteristics. Both equations can be standardized to give rate of change in standard deviation units if desired.

Of most importance to this analysis, however, is the prediction of total change resulting from the joint influence of variables at both levels. Computing this rate of change requires that the two partial derivatives be combined. To effect this combination, we calculate the total differential. The total differential predicts the magnitude of influence of changes in size (of both schools and districts) on school performance (which is always the dependent variable in these analyses), all else equal. Let us begin by explaining the partial derivatives. In the immediately subsequent section, however, we provide an explanation of and illustrate the use of the total differential, as it constitutes the most important interpretation of size effects jointly interaction with poverty.

Partial Derivatives. In Tables 3-15 we report two partial derivatives, one for each level of influence (school and district) separately. Partial derivatives give the rate of change in a dependent variable produced by focal variables (SPANSIZE and FREEPCT, in the present case), holding constant all other variables (i.e., BLACKPCT, MINORPCT, and S/RRATIO). Readers need to understand how they may use these additional equations. (Note 12) We will use the 8th grade composite statistics (Table 10) to illustrate our procedure, and we explain both the creation of partial derivatives and the calculation of the total differential. First, taking the partial derivative of Y with respect to SPANSIZE at the school level ("Y wrt 1" in Table 10) tells us that the rate of change in Y with respect to SCHOOL SPANSIZE, holding constant the other independent variables, is equal to:

$$f_{x1}(y) = [(-308.619)(\text{DISTRICT SPANSIZE})] - [(1.046)(\text{DISTRICT FREEPCT})]$$

Similarly, using the same outcome measure, taking the partial derivative of Y with respect to SPANSIZE at the district level tells us that the rate of change in Y with respect to DISTRICT SPANSIZE, holding constant the other independent variables, is equal to:

$$f_{x_2}'(y) = [(-308.619)(\text{SCHOOL SPANSIZE})] - [(4.304)(\text{SCHOOL FREEPCT})]$$

The first partial derivative enables us to see that, all else equal, if we increased the value of DISTRICT SPANSIZE by, say, one quarter standard deviation unit (.025 = .25 x .101), the predicted outcome measure would decrease by 7.7 points. Similarly, if DISTRICT FREEPCT were increased by one quarter standard deviation unit (4.4 = .25 x 17.5), the outcome measure would decrease by 4.6 points. These effects, of course, are additive, and changes of equal magnitude, but in the contrary directions, would yield no net effect.

The second partial derivative enables us to determine the effect on 8th grade composite scores of an increase or decrease in SCHOOL SPANSIZE and SCHOOL FREEPCT. A one quarter standard deviation unit increase in SCHOOL SPANSIZE (.031 = .25 x .124) yields a 9.6 point decrease in the outcome measure. A one quarter point standard deviation unit increase in SCHOOL FREEPCT (5.73 = .25 x 22.9) yields a 24.7 point decrease in the outcome measure.

The Total Differential

Information about the composite relationship between size and achievement is provided by the total differential. The total differential (dy) is the sum of the products of the partial derivatives and their differentials, dx₁ and dx₂, where dx₁ represents a change in SCHOOL SPANSIZE and dx₂ represents a change in DISTRICT SPANSIZE. The total differential, then, is the sum of the changes in measured achievement due to changes in SCHOOL SPANSIZE and DISTRICT SPANSIZE, contingent on SCHOOL FREEPCT and DISTRICT FREEPCT (all else equal):

$$dy = \{[f_{x_1}'(y)](dx_1)\} + \{[f_{x_2}'(y)](dx_2)\}$$

The values of dx₁ and dx₂ represent proportional changes (e.g., -.10 or +.10) in school or district size (SPANSIZE). To illustrate the calculation of the total differential, we computed hypothetical values of dx₁ and dx₂ tied to real-life values in the Georgia data set. We divided the SPANSIZE into the difference between SPANSIZE and the difference between the value of SPANSIZE for cases n + 1 and case n. That is, using the subsequent case in the data set as a reference point, we inferred rates change for school and district size in the subject case. This procedure produces arbitrary changes, but these arbitrary changes vary only within the range of variation that the Georgia school system exhibits.

In keeping with Dowling's (1980) admonition that differentials should be realistically small, we then eliminated cases with values for dx₁ or dx₂ greater than one-half standard deviation above or below their mean. (Note 13) The absolute value of dx₁ for all remaining cases was less than .068, and the absolute value of dx₂ was less than .026. We then randomly selected ten of the remaining schools for inclusion in Table 16.

Table 16
Total Differential: Illustrative Values for Randomly Selected Cases

Grade Eight Composite Scores

DISTRICT SPANSIZE	SCHOOL SPANSIZE	DISTRICT FREEPCT	SCHOOL FREEPCT	dx ₁	dx ₂	dy
.0829	.0835	79.47	71.79	-.047	-.010	8.47
.1562	.2187	73.38	74.90	-.019	-.009	6.08
.2285	.3427	20.21	0.90	-.066	-.005	6.63
.1541	.1527	24.34	27.10	.013	.018	-3.88
.1437	.2770	70.84	66.20	-.029	-.013	8.21
.1469	.1497	61.07	59.70	-.108	-.006	13.62
.1311	.1270	66.60	61.20	.005	.010	-3.56
.1825	.2120	29.76	19.50	.000	.005	-0.70
.0980	.0944	55.89	48.10	-.062	.013	2.46
.1566	.3060	47.39	55.30	.016	.016	-6.75

Notes.

Values of variables are given *uncentered*. Equations are derived from and computed with centered values.

Total differential computed as: $dy = \{[f_{x1}'(y)](dx_1)\} + \{[f_{x2}'(y)](dx_2)\}$

Values of partial differentials, dx₁ and dx₂ computed as follows (cases selected for $|dx_i| \leq 0.5\sigma$):

$$[(SPANSIZE_{case(n+1)}) - (SPANSIZE_{case(n)})] / (SPANSIZE_{case(n)})$$

The first four columns in Table 16 describe the focal variables (district and school size and subsidized meal rates). The fifth and sixth columns provide the (hypothetical) proportional changes in school size (dx₁) and district size (dx₂). The values of the total differential—the predicted change in each school's mean Composite Test score attributable to these composite changes in size—contingent on these proportional changes in school and district size appear in the column headed "dy" ("total differential").

Observe that Table 16 illustrates the inverse relationships between school performance (8th grade composite, in this case) and changes in SPANSIZE at both the school level and the district level. The first two cases, for instance, show a positive influence of joint school and district size in a uniformly impoverished school and district. Case seven shows the decline in similar circumstances of a joint increase in size. And case nine shows the somewhat more modest increase in test scores resulting from a joint reduction in school size and increase in district size.

Eighth and Eleventh Grade "Equity Effects"

Most people understand inequity in school finance. Affluent communities almost always enjoy better-funded schools, and improvements in financial equity would require that schools in impoverished communities be much better funded than they are. In other words, mitigating financial inequity requires that we break the link between poverty and school finance. Some educators (we among them) believe that no ethical principle justifies the privilege enjoyed by more affluent citizens in this regard. Why should the rich enjoy the best-funded schools? The rich commonly argue that it is their right, and the argument prevails.

Inequity in achievement presents much the same case. Which children, in general, enjoy the highest achievement? More affluent children do. Some observers, of course, believe that since the constructs "affluence" and "ability" correlate well, this state of affairs is actually very fair. The rich might well argue that inequity of outcomes in their favor is also their right. Others (we among them) note that—among affluent and impoverished people alike—a great range of abilities exists, and that in all adult occupations a similarly great range of abilities persists. On this view, the low achievement of impoverished children is not nearly so fair as it at first might seem (e.g., Gardner, 1983). In this view, public schooling can and should do much more to nurture the learning of impoverished students, in particular among all students. As with financial equity, equity in achievement means breaking—or at least substantially mitigating—the prevailing bond between SES and achievement. (Note 14)

Table 17
Multi-Level Georgia Equity Effects^a

Larger v. Smaller Schools and Districts with Grades 8 and 11

Composite					
		Grade 8		Grade 11	
		Districts			
		L^b	S	L	S
Schools	L	.76	.72	L .77	.74
	S	.63	.35	S .54	.16

Reading Comprehension (8)/ English (11)					
		Grade 8		Grade 11	
		Districts			
		L	S	L	S
Schools	L	.84	.74	L .69	.59
	S	.71	.36	S .28	.16

Mathematics

	Grade 8		Grade 11			
	Districts					
	L	S	L	S		
Schools	L	.71	.59	L	.72	.65
	S	.46	.29	S	.48	.25

	Science					
	Grade 8		Grade 11			
	Districts					
	L	S	L	S		
Schools	L	.82	.73	L	.73	.71
	S	.70	.37	S	.46	.27

Notes:

a) Variance (R^2) in school performance attributable to school-level subsidized meal rates.

b) L = Larger half; S = Smaller half.

Table 17 gives the variance in achievement associated with SES in four groups by the medians of district size and school size (2 grades and 4 tests). Within each panel, by grade level, we report the observed variances proceeding left to right and top to bottom in each of the 8 contrasts for: (1) large schools in large districts, (2) large schools in small districts, (3) small schools in large districts, and (4) small schools in small districts.

In each of these 8 (2 grade levels by 4 tests) four-way contrasts, large schools in large districts show the highest proportion of variance in achievement associated with SES: between 71% and 84%, whereas the lowest proportion of variance is exhibited among small schools in small districts: between 16% and 27%. Moreover, the order of declining variance follows an identical pattern in each of the 8 contrasts: large-large, large-small, small-large, and small-small. In 6 of 8 cases, the largest magnitude of decline within the evident sequence (large-large, large-small, etc.) of decreasing variance comes in the change from small schools in large districts to small schools in small districts.

In other words, Table 17 suggests that the predicted equity effect of reducing district size but not school size would be practically significant; the predicted equity effect of reducing school size but not district size would also be practically significant and perhaps somewhat larger; and the combined strategy of reducing both school and district size would be predicted to yield substantial equity and excellence effects (given the previous multi-level regression analyses).

Some rural states (e.g., Montana; see Howley 1999b) structure their school systems in just this way. That is, such systems have chosen to sustain small schools within small districts. The Montana system doubtless has plenty of room for "improvement," but on the terms of accountability (and the value of more equal outcomes), Montana is an exemplar. Please note that Montana has a substantial American Indian population (13%), whose children also attend small, predominantly public, schools and districts.

In rural areas, the phenomena of school and district size seem mutually dependent; larger rural schools often prevail in larger rural districts (e.g., as in West Virginia; see

Howley, 1996). District "reorganization" has often been a first step toward eliminating small schools (DeYoung, 1995; Peshkin, 1982). This strategy would be predictably harmful to the achievement of students in impoverished rural communities. In the southeast US a single high school now often serves entire rural counties, covering large geographical areas.

The situation in urban areas is equally bad, though in somewhat different ways. The huge big-city districts were created, not just to improve schools, but to destroy a resource (school jobs) that could be controlled by ward politics. Usually portrayed as a "progressive" change, an important motive of city fathers was to wrest power back from the hands of working-class urban communities (e.g., Tyack, 1974; Erie, 1988). Today, most urban districts are nightmares and wildernesses of bureaucracy and outright fear (e.g., Devine, 1996). Jobs are as much a political issue as ever in many large cities, but the power to dispense them has shifted to the nexus between political regimes and school bureaucracies, with the bureaucracy often in the better position. No wonder so many thoughtful educators champion the re-establishment of smaller schools in cities (e.g., Meier, 1995; Klonsky, 1995).

Difficult as it is, in both rural and urban locales, to defend or re-establish small schools, that task leaves the structural challenge incomplete. Seldom are reductions in district size—especially in the case of large city districts—seriously considered.

Our principal "clear and simple" recommendation therefore is to suggest the wisdom, of reorganizing districts that are now far too large. Policy makers should start imagining ways to re-create districts that are everywhere sufficiently small to respond well to students, families, and (especially) communities. One way to enable this decision making might be for communities to enjoy the right to charter public school districts as well as public schools (and, naturally, to receive the requisite state-level support to succeed). The policy issues are surely difficult, but no more difficult than those that have already led to the counterproductive structuring that presently prevails. To do nothing or little leaves the burden of coping with the enormity to impoverished students, families, and communities—exactly where it currently rests.

Misuse of the findings

Our findings cannot be interpreted to warrant the construction of huge schools, however, even for relatively comfortable communities; in general, we advise an upper limit of about 250 students per grade for 9-12 high schools and about 100 students per grade for elementary schools—and these rule-of-thumb upper limits apply to communities where the poverty rate is zero (Howley, 1997; but see Irmsher, 1997, and Raywid, 1999, for quite similar recommendations based on recent reviews of the literature).

Recently we learned that our research was being used to help justify construction of a school in a semi-rural area of an eastern state proposed to house 2,000 elementary students in grades 3-6. In view of extant and easily accessible research syntheses such as those by Irmsher and Raywid, proposals to create schools of this size—particularly elementary schools—are, we believe, capricious and professionally irresponsible.

We are unhappy (but not surprised) to learn that our work has been deployed to support such proposals; but we also understand the role that bad state-level policy plays in shaping such decisions as this (see Purdy, 1997, for a clear example in a rural state where the state influence is heavy-handed). The administration in this district experienced considerable angst when community members there contacted us and we voiced our objections to the misuse of our research publicly. In fact, however, we are

used to being contacted by community members resisting such efforts and equally used to not hearing from members of our own profession as they make construction plans. Despite uproar in the community and defeat of the bond issue, plans for the mega-school (to be organized in "houses") apparently continue. The superintendent in this case has reportedly vowed revenge on the interfering outside researchers! We regret the angst that emerges in these situations, but we believe the present study provides evidence to support our evolving position on the issues.

Conclusion

Small size is good for the performance of impoverished schools, but it now seems as well that small district size is also good for the performance of such schools in Georgia, where district size, in single-level analyses, had revealed no influence. Because of the consistency of school-level findings in previous analyses, we strongly suspect that the Georgia findings characterize relationships in most other states. This claim can, of course, only be evaluated by additional replications, and we hope other researchers will see merit in such work.

The equity effects reported here, however, extend the evidence of the previous single-level studies to the interaction of school and district size. Larger schools in larger districts seem to propagate inequality of outcomes by comparison to smaller schools and smaller districts. In fact, smaller schools in larger districts demonstrate a useful equity effect, as well. For large schools in smaller districts, however, the improvements in equity might be so slight as to be called "negligible."

The equity effects are so striking, and appear so instrumental in association with the "excellence" effects of small size in impoverished communities, that further investigation into this mitigating influence would seem crucial. How does the principle evident in the findings apply to individual students? In what settings? To what extent? What structural features of small size enable such an effect? How do impoverished students fare in schools that are, overall, rather affluent? Is an overall upper limit to school size and district size worth establishing by policy? How should such upper limits be set? What policies can succeed in recreating smaller districts in big cities and the rural southeast?

These are interesting and important questions, we think, but the conclusions of this study would seem to require rather wide debate and reconsideration of the size issue, across the spectrum of poverty and wealth, and not just in the case of impoverished communities. We note that America's elite sends its children to Andover and Exeter and other such fine high schools, where enrollments seldom exceed 1,500. What do they know that the rest of us have yet to learn, we wonder?

Notes

This work is based on research funded by the Rural School and Community Trust. It does not, however, represent the opinions or positions of the Rural Trust. We are grateful for the support; the errors and opinions are our own.

The two authors are equal contributors to the work reported here.

1. Unlike the other states, Montana has retained many small schools, and this historic decision is a likely cause for the weak interaction effects. Bickel (1999b) also reported no interaction effect among the 132 Texas schools that house all students

- in grades K-12 (cf. Franklin & Glascock, 1998).
2. Magnitude of the relationship was measured as the proportion of variance in achievement associated with SES.
 3. See the Appendix for a discussion of the problem that intraclass correlation poses to the use of ordinary least squares regression. The Appendix describes the conditions needed to use OLS in multi-level analysis and shows that our data set meets these conditions.
 4. This logic, of course, is also supported by the findings previously reported for the single-level four-state analyses, in which reported effects are strongest at grade 8 or 9, and always weaker at grades 11 or 12. On the basis of past experience, then, we would have reason to suspect similar results.
 5. We report statistical significance levels as a gauge to practical significance. Because the data set includes practically all schools in Georgia, the relationships that emerge are those that prevail, and, we maintain, should not be considered as subject to sampling error.
 6. Directionality of the influence is given in parentheses following the variable name. The effect of centering is not reflected in Tables 1 and 2.
 7. These findings are conceptually consistent with previously reported school-level analyses, which found that, among impoverished communities, smaller schools reduced the achievement costs of poverty and that large ones magnified such costs; but the converse was true as well, in those cases: Among affluent communities, smaller schools increased the achievement costs of affluence and larger ones reduced such costs.
 8. "Mean achievement costs" represent declines in predicted achievement. Therefore, another way to put this interactive relationship is this: (1) as poverty and district size continuously increase, predicted school performance continuously declines; and (2) as poverty decreases and district size decreases, predicted performance also continuously declines.
 9. We might also observe that other cross-level interactions appear significantly in the equations reported in several of these Tables: Table 11 (math: FREEPCT), Table 13 (social studies: SPANSIZE), and Table 14 (science: BLACKPCT). Cross-level structural influences are weak at the 11th grade but still evident.
 10. Robert Moses's "Algebra Project" construes algebra as the course that governs access to the academic track in life; failing algebra, or never taking it in the first place, marks one as academically inept.
 11. A "derivative" can be understood as the calculus tool for determining the "slope" of a curved line (which, in geometrical terms, is the tangent of the curve at a given point). The slope of such a line is constantly changing (just as the effects of school or district size, or their joint effects, constantly change with respect to poverty levels), and the derivative provides the formula for calculating this changing slope. To find this changing rate, one "takes the derivative" of the formula that describes the line. A partial derivative holds one variable constant during differentiation (the process of "taking the derivative") so that the influence of that variable can be subsequently evaluated. This process of "holding an influence constant" is similar to calculating a partial correlation coefficient.
 12. Consult Howley (1996) for a complete description of the derivation of partial derivatives in the single-level analyses.
 13. Dowling's counsel is important because we are dealing, in using calculus techniques that estimate changing rates, with how these rates of change at the margin (i.e., the usual addition or loss of a few students) under normal conditions,

and not, in fact, in such catastrophic alterations as are produced by consolidations of two or more schools (where size may well increase by hundreds of students). Calculus is the mathematics of smooth curves and not of disruption and disjunction.

14. In practical terms, one is unlikely to break the bond completely, because the negative effects of poverty can be eliminated only when a society finds them intolerable and actively cultivates the well-being of the poor. Even in the current economic boom, however, such a realization has not overtaken the US, and in general, the gap between the affluent and the impoverished is growing ever wider here. Also, some observers balk when they realize that breaking the bond must apply not just to the poor, but to the affluent as well.

References

- Alspaugh, J. (1995, October). *A comparison of four enrollment groups of K-8 and K-12 Missouri rural school districts*. Paper presented at the annual meeting of the National Rural Education Association, Salt Lake City, UT. (ERIC Document Reproduction Service No. ED 389 501)
- Barcikowski, R. (1981) Statistical power with the group mean as the unit of analysis. *Journal of Educational Statistics*, 6(3), 267-285.
- Barker, R., & Gump, P. (1964). *Big school, small school*. Stanford, CA: Stanford University Press.
- Barr, R., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.
- Betts, J. (1996). Is there a link between school inputs and earnings? Fresh scrutiny on an old literature. In Burtless, G. *Does money matter?* Washington, DC: Brookings.
- Bickel, R. (1999a). *School size, socioeconomic status, and achievement: A Georgia replication of inequity in education*. Randolph, VT: Rural Challenge Policy Program.
- Bickel, R. (1999b). *School size, socioeconomic status, and achievement: A Texas replication of inequity in education*. Randolph, VT: Rural Challenge Policy Program.
- Bidwell, C., & Kasarda, J. (1975). School district organization and student achievement. *American Sociological Review*, 40(1), 55-70.
- Boyd, L., & Iversen, G. (1979). *Contextual analysis: Concepts and statistical techniques*. Belmont, CA: Wadsworth Publishing Co.
- Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.
- Bryk, A., & Raudenbush, S. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Burtless, G. (Ed.). (1966). *Does Money Matter*: Washington, DC: Brookings.
- Caldas, S. (1993). Re-examination of input and process factor effects on public school achievement. *Journal of Educational Research*, 86, 206-214.

- Clones, D. (1998). *Development report card for the states* (12th ed.). Washington, DC: Corporation for Enterprise Development.
- Conant, J. (1959). *The American high school today*. New York: McGraw-Hill.
- Cronbach, L. (1987) Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102(3), 414-417
- Devine, J. (1996). *Maximum security*. Chicago: University of Chicago Press.
- Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives*, 6(1) [On-line serial]. Available World Wide Web: <http://olam.ed.asu.edu/epaa/v5n12.html>
- Dowling, E. (1980). *Mathematics for economists*. New York: McGraw-Hill.
- Erie, S. (1988). *Rainbow's end: Irish Americans and the dilemmas of urban machine politics*. Berkeley, CA: University of California Press.
- Farkas, G. (1996). *Human capital or cultural capital? Ethnicity and poverty groups in an urban school district*. Hawthorne, NY: de Gruyter.
- Fowler, W. (1995). School Size and Student Outcomes. *Advances in educational productivity*, 5, 3- 26.
- Fowler, W. (1991). School size, characteristics, and outcomes. *Educational evaluation and policy analysis*, 13, 189-202.
- Franklin, B., & Glascock, C. (1998). The relationship between grade configuration and student performance in rural schools. *Journal of Research in Rural Education*, 14(3), 149-153.
- Friedkin, N., & Necochea, J. (1988). School system size and performance: A contingency perspective. *Educational Evaluation and Policy Analysis*, 10(3), 237-249.
- Gamoran, A., & Dreeben, R. (1986). Coupling and control in educational organizations. *Administrative Science Quarterly*, 31: 612-632.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Georgia Department of Education. (1999). *A message from the state superintendent of schools: Welcome to the 1997-1998 Georgia public education report card*. Atlanta, GA: Author. (Available World Wide Web: <http://168.31.216.185/RCP/LindaLtr.html>)
- Goldstein, H. (1995). *Multilevel statistical models*. London: Edward Arnold.
- Gujarati, D. (1995). *Basic econometrics* (3rd ed.). New York: McGraw-Hill.
- Guthrie, J. (1979). Organizational scale and school success. *Educational Evaluation and Policy Analysis*, 1, 17-27.

- Haller, E. (1992). High school size and student indiscipline: Another aspect of the school consolidation issue. *Educational Evaluation and Policy Analysis, 14*, 145-156.
- Haller, E., Monk, D., & Tien, L. (1993). Small schools and higher order thinking skills. *Journal of Research in Rural Education, 9*, 66-73.
- Hanushek, E. (1998). Conclusions and Controversies about the Effectiveness of School Resources. Federal Reserve Bank of New York Economic Policy Review, 4: 11-28.
- Hanushek, E. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*: 141-164.
- Hemmings, B. (1996). A longitudinal study of Australian senior secondary school achievement. *Issues in Educational Research, 6*, 13-37.
- Henderson, R., & Raywid, M. (1996). A "small" revolution in New York City. *Journal of Negro Education, 63*, 28-45.
- Howley, C. (1999a). *The Matthew Project: State report for Montana*. Randolph, VT: Rural Challenge Policy Program.
- Howley, C. (1999b). *The Matthew Project: State report for Ohio*. Randolph, VT: Rural Challenge Policy Program.
- Howley, C. (1997). Dumbing down by sizing up: Why smaller schools make more sense—if you want to affect student outcomes. *The School Administrator, 54*(9), 24-26, 28, 30.
- Howley, C. (1996). The Matthew Principle: A West Virginia Replication? *Education Policy Analysis Archives, 3*(1), (Entire Issue.) [Available online: <http://epaa.asu.edu/epaa/v3n18.html>]
- Howley, C. (1989). *What is the effect of small-scale schooling on student achievement?* (Digest EDO-RC-89-6). Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Document Reproduction Service No. ED 308 062)
- Howley, C. (1989). Synthesis of the effects of school and district size: What research says about achievement in small schools and school districts. *Journal of Rural and Small Schools, 4*, 2-12.
- Howley, C., & Bickel, R. (1999). *The Matthew Project: National report*. Randolph, VT: Rural Challenge Policy Program.
- Huang, G., & Howley, C. (1993). Mitigating disadvantage: Effects of small-scale schooling on student achievement in Alaska. *Journal of Research in Rural Education, 9*(3), 137-149.
- Irmsher, K. (1997). *School size* (ERIC Digest). Eugene, OR: ERIC Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. ED 414 615)
- Iversen, G. (1991). *Contextual analysis*. Newbury Park, CA: Sage.

- Khattari, N., Riley, K., & Kane, M. (1997). Students at risk in poor rural areas: A review of research. *Journal of Research in Rural Education*, 13, 79- 100
- Klonsky, M. (1995). *Small schools: The numbers tell the story: A review of recent research and experiences*. Chicago, IL: Small Schools Workshop. (ERIC Document Reproduction Service No. ED 386 517)
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
- Lamdin, D. (1995). Testing for the effect of school size within a district. *Educational Economics*, 3, 33-42.
- Lee, V. and Smith, J. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis*, 19(3), 205-227.
- McDill, E., Natriello, G., & Pallas, A. (1986). A population at risk: Potential consequences for tougher school standards for student dropouts. In G. Natriello (Ed.), *school dropouts: Patterns and policies*. New York: Teachers College Press.
- Meier, D. (1995). *The power of their ideas: Lessons for America from a small school in Harlem*. Boston: Beacon Press.
- Meyer, J., Scott, W., & Strang, D. (1987). Centralization, fragmentation, and school district complexity. *Administrative Science Quarterly*, 32(2), 186-201.
- Mik, M., & Flynn, M. (1996). School size and academic achievement in the HSC examination: Is there a relationship? *Issues in Educational Research*, 6, 57- 78.
- Peshkin, A. (1982). *The imperfect union: School consolidation and community conflict*. Chicago: University of Chicago Press.
- Prichard Committee for Academic Excellence. (1990). *The path to a larger life: Creating Kentucky's educational future*. Lexington, KY: University of Kentucky Press. (ERIC Document Reproduction Service No. ED 323 055)
- Purdy, D. (1997). An economical, thorough, and efficient school system: The West Virginia School Building Authority "economy of scale" numbers. *Journal of Research in Rural Education*, 13, 70-82.
- Raywid, M. (1999). *Current literature on small schools* (EDO-RC-98-8). Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Document Reproduction Service No. ED 425 049)
- Raywid, M. (1996). *Taking stock: The movement to create mini-schools*. New York: ERIC Clearinghouse on Urban Education. (ERIC Clearinghouse Accession No. UD 031 001)
- Riley, R. (1999, October 13). *Schools as centers of community* (Remarks as prepared for delivery by Secretary of Education Richard W. Riley to the American Institute of Architects). Washington, DC: US Department of Education. (Available World Wide

Web: <http://www.ed.gov/Speeches/10-1999/991013.html>)

Riordan, C. (1997). *Equality and achievement*. New York: Longman.

Rivkin, S., Hanushek, E., & Kain, J. (1998). *Teachers, schools, and academic achievement* (Working Paper Number 6691). Washington, DC: National Bureau of Economic Research.

Rossmiller, R. (1987). Achieving equity and effectiveness in schooling. *Journal of Educational Finance*, 4, 561-577.

Rural Trust. (1999, November). Riley recognizes community role in school design and use. *Rural Policy Matters*, p. 2.

Silva, C., & Moses, R. (1990). The algebra project: Making middle school mathematics count. *Journal of Negro Education*, 59(3), 375-91.

Singer, J. (1987) An intraclass correlation model for analyzing multilevel data. *Journal of Experimental Psychology*, 18(4), 219-228.

Smith, D., & DeYoung, A. (1988). Big school vs. small school: Conceptual, empirical, and political perspectives on the re-emerging debate. *Journal of Rural and Small Schools*, 2, 2-11.

Stevens, N., & Peltier, G. (1994). A review of research on small-school participation in extracurricular activities. *Journal of Research in Rural Education*, 10, 116-120.

Stevenson, K. (1996). Elementary school capacity: What size is the right size? *The Educational Facility Planner*, 33, 10-14.

Stiefel, L., Berne, R., Iatarola, P., & Fruchter, N. (2000). High school size: Effects on budgets and performance in New York City. *Educational Evaluation and Policy Analysis*, 22(1), 27-39.

Thiétart, R.A., & Forgues, B. (1995). Chaos theory and organization. *Organization Science*, 6(1), 19-31.

Tyack, D. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.

Walberg, H., & Walberg, H. (1994). Losing local control of schools. *Educational Researcher*, 23, 19- 26.

Wyatt, T. (1996). School effectiveness research: Dead end, damp squib, or smoldering fuse? *Issues in Educational Research*, 6, 79-112.

About the Authors

Robert Bickel

Marshall University

Email: bickel@marshall.edu

Robert Bickel is a Professor of Advanced Educational Studies at Marshall University. His recent research is concerned with school size as a variable which moderates the relationship between social class and measured achievement, evaluation of early childhood interventions, and with contextual factors which occasion the at-risk designation.

Craig Howley

Appalachia Educational Laboratory, Inc.
Ohio University
75619 Lively Ridge Road
Albany, OH 45710
Phone: 740.698.0309
Email: howleyc@ael.org

Craig Howley is an education writer and researcher based in Albany, Ohio, and affiliated part-time with Ohio University (Athens, OH) and AEL, Inc. (Charleston, WV). Ongoing scholarly projects in which he is a partner include the following: research on school and district size, rural school busing, and three book projects: an examination of small rural high schools (for Appalachian Educational Laboratory), an extended interpretation of developmentalism as a school ideology (with Aimee Howley) and a textbook (with A. Howley and Ohio University colleagues) on school administration.

Appendix

Ordinary Least Squares Regression and the Problem of Intraclass Correlations

One of the assumptions of ordinary least squares estimators is that residuals are not correlated. However, in a multi-level analysis this assumption may be erroneous. The reason is that first-level observations are located within the groups that constitute the second level of analysis. Grouping of first-level observations (schools) into districts may well mean that schools within a district are more like each other than they are like schools in other districts. The consequence is intraclass correlation, or covariance among residuals for schools in the same district (see Kreft & de Leeuw, 1998, pp. 9-10).

This observation yields the primary objection to traditional contextual models such as ours. Through uncritical use of ordinary least squares, the magnitude of standard errors of regression coefficients may be underestimated and alpha levels artificially inflated (Goldstein, 1995). The observation holds even though ordinary least squares estimators remain unbiased (Barcikowski, 1981).

In the present study, intraclass correlations, which vary by outcome measure and grade level, range in magnitude from .048 to .101. The number of groups or districts is 158 for the 8th grade and 155 for the eleventh grade. With 367 schools reporting 8th grade test scores, and 298 reporting eleventh grade scores, the relative number of second-level observations is large, indeed (Goldstein, 1995).

We conclude that intraclass correlation is a negligible problem. Given this confluence of circumstances—small intraclass correlations and large numbers of districts relative to the number of schools—ordinary least squares will yield estimates which are unbiased and will provide such estimates with very little inflation of

regression coefficient variances (Singer, 1987). Furthermore, using a procedure presented by Singer (1987), we have calculated the remaining modest inflation of regression coefficient variances, standard errors, and resulting t-values. We compensated for this statistical artifact when running tests of significance, reducing the magnitude of the affected statistics by the amount they are inflated due to intraclass correlation.

Copyright 2000 by the *Education Policy Analysis Archives*

The World Wide Web address for the *Education Policy Analysis Archives* is epaa.asu.edu

General questions about appropriateness of topics or particular articles may be addressed to the Editor, [Gene V Glass](mailto:glass@asu.edu), glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-0211. (602-965-9644). The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu .

EPAA Editorial Board

[Michael W. Apple](#)
University of Wisconsin

[John Covalleskie](#)
Northern Michigan University

[Sherman Dorn](#)
University of South Florida

[Richard Garlikov](#)
hmwkhel@scott.net

[Alison I. Griffith](#)
York University

[Ernest R. House](#)
University of Colorado

[Craig B. Howley](#)
Appalachia Educational Laboratory

[Daniel Kallós](#)
Umeå University

[Thomas Mauhs-Pugh](#)
Green Mountain College

[William McInerney](#)
Purdue University

[Les McLean](#)
University of Toronto

[Anne L. Pemberton](#)
apembert@pen.k12.va.us

[Richard C. Richardson](#)
New York University

[Greg Camilli](#)
Rutgers University

[Alan Davis](#)
University of Colorado, Denver

[Mark E. Fetler](#)
California Commission on Teacher Credentialing

[Thomas F. Green](#)
Syracuse University

[Arlen Gullickson](#)
Western Michigan University

[Aimee Howley](#)
Ohio University

[William Hunter](#)
University of Calgary

[Benjamin Levin](#)
University of Manitoba

[Dewayne Matthews](#)
Western Interstate Commission for Higher Education

[Mary McKeown-Moak](#)
MGT of America (Austin, TX)

[Susan Bobbitt Nolen](#)
University of Washington

[Hugh G. Petrie](#)
SUNY Buffalo

[Anthony G. Rud Jr.](#)
Purdue University

Dennis Sayers
Ann Leavenworth Center
for Accelerated Learning

Michael Scriven
scriven@aol.com

Robert Stonehill
U.S. Department of Education

Jay D. Scribner
University of Texas at Austin

Robert E. Stake
University of Illinois—UC

David D. Williams
Brigham Young University

EPAA Spanish Language Editorial Board

Associate Editor for Spanish Language
Roberto Rodríguez Gómez
Universidad Nacional Autónoma de México

roberto@servidor.unam.mx

Adrián Acosta (México)
Universidad de Guadalajara
adrianacosta@compuserve.com

Teresa Bracho (México)
Centro de Investigación y Docencia
Económica-CIDE
bracho disl.cide.mx

Ursula Casanova (U.S.A.)
Arizona State University
casanova@asu.edu

Erwin Epstein (U.S.A.)
Loyola University of Chicago
Eepstein@luc.edu

Rollin Kent (México)
Departamento de Investigación
Educativa-DIE/CINVESTAV
rkent@gemtel.com.mx
kentr@data.net.mx

Javier Mendoza Rojas (México)
Universidad Nacional Autónoma de
México
javiermr@servidor.unam.mx

Humberto Muñoz García (México)
Universidad Nacional Autónoma de
México
humberto@servidor.unam.mx

Daniel Schugurensky
(Argentina-Canadá)
OISE/UT, Canada
dschugurensky@oise.utoronto.ca

J. Félix Angulo Rasco (Spain)
Universidad de Cádiz
felix.angulo@uca.es

Alejandro Canales (México)
Universidad Nacional Autónoma de
México
canalesa@servidor.unam.mx

José Contreras Domingo
Universitat de Barcelona
Jose.Contreras@doe.d5.ub.es

Josué González (U.S.A.)
Arizona State University
josue@asu.edu

María Beatriz Luce (Brazil)
Universidad Federal de Rio Grande do
Sul-UFRGS
luceb@orion.ufrgs.br

Marcela Mollis (Argentina)
Universidad de Buenos Aires
mmollis@filo.uba.ar

Angel Ignacio Pérez Gómez (Spain)
Universidad de Málaga
aiperez@uma.es

Simon Schwartzman (Brazil)
Fundação Instituto Brasileiro e Geografia
e Estatística
simon@openlink.com.br

Jurjo Torres Santomé (Spain)
Universidad de A Coruña
jurjo@udc.es

Carlos Alberto Torres (U.S.A.)
University of California, Los Angeles
torres@gseisucla.edu
