DOCUMENT RESUME

ED 433 174 RC 022 082

AUTHOR Howley, Craig B.; Bickel, Robert
TITLE The Matthew Project: National Report.
INSTITUTION Ohio State Univ., Columbus.; Marshall Univ., Huntington, WV.; Appalachia Educational Lab., Charleston, WV.
SPONS AGENCY Rural Challenge Policy Program, Randolph, VT.
PUB DATE 1999-08-25
NOTE 29p.; For related reports, see RC 022 081, RC 022 083, and RC 022 087-88.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Economically Disadvantaged; Elementary Secondary Education; *School District Size; *School Size; *Small Schools; *Socioeconomic Influences; *Socioeconomic Status
IDENTIFIERS Georgia; Montana; Ohio; Small School Districts; Texas

ABSTRACT

Previous studies found that the small size of schools or school districts mitigated the negative influence of poverty on academic achievement in California, Alaska, and West Virginia. The Matthew Project extends this research in four additional states selected to provide varied settings: Ohio, Georgia, Texas, and Montana. Data from each state were used in regression equations that predict overall school or district achievement from measures of size, socioeconomic status (SES), and the product of size and SES. These equations illuminate possible "excellence effects" of size by showing which communities (based on SES-level) may benefit or lose from increases in school or district size. Equity effects of size on achievement were also tested by computing the correlation between SES and achievement in groups of larger and smaller schools and districts. Strong evidence of an interaction effect of school size was found in Ohio, Georgia, and Texas, such that academic achievement benefited from smaller schools in more impoverished communities and from larger schools in more affluent communities. A weaker interactive effect was found in Montana, which maintains many small schools. Across all four states, a strong equity effect was found at all grade levels, whereby small size reduced the negative influence of poverty on school and district performance. Strong evidence of an interaction effect of district size was found only in Ohio. The Matthew Project studies indicate that a one-best, everywhere "optimal," school size is a figment. Four policy questions related to school and district size are discussed. (Contains 15 references.) (SV)

The Matthew Project:
National Report

Craig B. Howley
Ohio University and Appalachia Educational Laboratory
75619 Lively Ridge Road, Albany, OH 45710
howleyc@ael.org

Robert Bickel
Marshall University
College of Education and Human Services
400 Hal Greer Boulevard, Huntington, WV 25725-2440
bickel@marshall.edu

August 25, 1999

The Matthew Project: National Report

EXECUTIVE SUMMARY

The Matthew Project, with support from the Rural Challenge Policy Program, investigated the possible academic excellence and equity effects of school and district size in Montana, Georgia, Ohio, and Texas (Bickel, 1999a, 1999b; Howley, 1999a, 1999b). Previous studies (Friedkin & Necochea, 1988; Howley, 1995, 1996; Huang & Howley, 1993) had reported that the effects of size depended on accounting for the interaction of size and socioeconomic status (SES). The Matthew Project studies build on that line of inquiry.

Method

We constructed equations that predict overall school or district test scores from measures of size, socioeconomic status, and the product of size and socioeconomic status. These (regression) equations provide a view of the possible ***excellence effects*** of size because they show which SES levels are likely to benefit from, or conversely, to be harmed from increases in school or district size, and to what extent.

We also tested the ***equity effects*** of size (on academic achievement) by dividing relevant groups of schools and districts into two equal groups at the median of size. Then we computed the correlation between SES and achievement. Weaker correlations indicate that SES accounts for less of the variability in achievement, and so indicate a weakening of the bond between these two qualities. Equity in achievement depends on the weakening of this bond; for instance, on *disrupting* or *mitigating* the negative influence of poverty on achievement. This goal reflects that view, for instance, that achievement should be more strongly influenced by such qualities as effort, adequacy of educational funding, and fairly distributed opportunities to learn.

Findings

- Strong evidence of an interaction effect of ***school size*** exists in Georgia, Ohio, and Texas. In these three states, ***the poorer the community the smaller should schools be in order to maximize school performance as measured by standardized tests***. The interaction effect, of course, also suggests that performance in more affluent communities benefits from larger schools.

- Across all four states, on virtually all measures and at all grade levels, a strong ***equity effect of small size*** (schools and districts) exists. This effect reduces the negative influence of poverty on school and district performance by as much as 70% and seldom, in our analyses,

by less than 20%.

- Evidence for an interaction effect of school size exists in ***Montana***, but it is much weaker than in the other three states. Unlike the other states, Montana maintains many small schools, at all grade level.

- Strong evidence of an interaction effect of ***district size*** exists only in our Ohio analyses. In Ohio, larger districts differentially benefit performance for more affluent students--the more affluent the community, the stronger the positive benefits.

Conclusions

1. In view of previous findings in Alaska, California, and West Virginia, the new findings for Georgia, Montana, Ohio, and Texas suggest that an interaction effect of school size is prevalent in the US. This prevalence seems, however, to be sharply limited in states, such as Montana, that maintain smaller schools.

2. A related and even stronger finding concerns the apparent capacity of small schools to break the bond between SES and school performance. In small schools, the relationship between SES and school achievement is substantially weaker . The consistency of this finding is strong evidence for the widespread prevalence of this effect.

3. An interaction effect for district size may exist in some states; the Matthew Project found such an effect in Ohio. Previous research found such an effect in West Virginia and California. Some states may maintain districts too small or too large to maximize the aggregate achievement of the sorts of students they serve.

Policy Questions

Good policy should aim to distribute resources in such a way that students of various SES levels are benefitted to the maximum. Such arrangements would, in effect, cultivate a degree of excellence by providing conditions that appear to maximize academic achievement under varying circumstances.

Citizens and policy makers in the states might consider four questions related to the issue of district and school size. Answering these questions *appropriately* in different states has the potential to improve education for all students, but particularly for students from impoverished communities. Though not answered very easily, the important questions are stated quite simply:

1. Should states set an absolute upper limit on the sizes of districts and schools? If the answer is "yes," then what should these caps be for elementary, middle, and high schools and for districts?

2. In states where an interaction effect exists, which districts and schools violate these limits? What should be done in these cases?

3. In states where an interaction effect exists, how should such upper limits vary by school and district SES? Should *any* lower limits prevail? What would they be?

4. What policies might successfully promote the re-scaling ("restructuring") of state systems of schooling as needed the better to serve all the communities whose children attend them?

Equity and excellence cannot be pursued separately; one is pursued single-mindedly only at the expense of the other. The equity effect disclosed by our analyses suggests that all schools should be *smaller*. But that inference from the analyses does not say how small schools should by, or, more particularly, how small *for whom*. The interaction findings provide some guidance. They suggest that the smallest schools *will be productively reserved* for the most impoverished communities. The good news is that, in order to maximize student achievement, schools in affluent communities can be substantially larger. The point is that equity and excellence can be maximized simultaneously, and, further, maximizing one independent of the other is destructive of the health of the educational system, and, ultimately, of national well-being.

This report suggests a ***rule of thumb for determining upper limits*** of school size, with 1,000 as the usual upper limit for a 9-12 high school serving a very affluent community. Simple interpolation will give equivalent sizes for elementary and middle schools, or schools of varying grade-span configurations. But the upper limit could be 1,500. Our Montana report, though, shows that the lower extreme can be successfully be much smaller--high schools with fewer than 100 students and elementary schools with even as few as 10 students. And our analyses suggest that schools for very impoverished communities *should* be much smaller.

So far, and unfortunately for impoverished communities, state education agencies have concentrated on determining the ***lower limits*** of school size. The fixation on lower limits of school size is a legacy of the early days of building the national US system of schools, when the objective was to close small one-teacher schools as widely and quickly as possible and to create larger and more centralized school districts. The strange fact is that this thinking persists. The time may be coming, nonetheless, for this fixation to weaken.

The Matthew Project: National Report

Research on the consequences of variability in school size has a long history. As with so many variables in educational research, empirical investigations of school size effects have, over the years, yielded conflicting results. This state of affairs has led some researchers to treat school size as a control variable which they are obliged to employ, but which is otherwise uninteresting.

Recent research, however, has linked school size, and also district size, to both effectiveness and equity in a new and interesting way: as size increases, some have found, the mean achievement costs for schools with less-advantaged students become more burdensome. The first reports of this interesting finding and its educational policy implications were based on research using data from California (Friedkin & Necochea, 1988), Alaska (Huang & Howley, 1993), and West Virginia (Howley, 1995, 1996). In an effort to determine if results from these three very different states can be generalized to other settings, we have replicated the research using data sets from four additional states chosen for their diverse demographic, geographic, political-economic, and educational circumstances: Georgia, Montana, Ohio, and Texas.

Background

Poverty figures as the chief and most prevalent threat to normal academic accomplishment among individuals. If your family is poor, your own odds of succeeding in school lengthen. Your odds are longer still if you attend school with children from many other

poor families, which is a likelihood in the U.S., since schools are segregated by social class[1]. In any case, it is certain that affluent communities enjoy decent schools and high-minded pedagogy, whereas impoverished communities continue to "enjoy" shabby schools and a pedagogy of expedience (that is, schooling that is primarily custodial). In fact, one might say that as the threats increase among increasingly impoverished communities, the local resources to counter them diminish *simultaneously*. It seems an especially vicious arrangement.

If, however, some quite ordinary and easily appreciated feature of schooling could be so deployed as to resolve this dilemma more favorably for the children of society's least privileged members, we should applaud it and move to deploy it as seemed advisable. The Matthew Project has pursued a promising line of inquiry relevant to such a hope.

This line of inquiry tests the "interaction hypothesis" of school and district size. The interaction hypothesis expresses the possibility that the degree (i.e., strength or weakness) and directionality (positive or negative) of the relationship of size to achievement is contingent on community SES. That is, no one size is "best" or "optimal," because the effects of size hypothetically vary among communities with differing levels of SES. The interaction hypothesis suggests that in some places the relationship could be negative and in some places positive; and that, in some places it could be weak, and in some places strong. Further, as a formal and testable hypothesis, it says that this variation could be *systematically associated* with changes some other condition. What might such a condition be? Socioeconomic status (SES) is well

---

[1]San Francisco recently adopted a plan to integrate its schools on a socioeconomic instead of a racial basis. "Ironically, the fact that economic segregation was never found unconstitutional means that voluntary measures addressed to class are constitutionally permissible" (Kahlenberg, 1999, p. 30-52).

known to be the strongest single influence on student achievement, and so it is a logical choice for our "contingent condition."

In the Matthew Project studies, we use school and district performance on state-mandated standardized tests as the measure of achievement. In some states these tests are norm-referenced and in some states they are criterion-referenced (for instance, some states require students to "pass" proficiency tests, and the percent passing in a school or district becomes a gauge of accountability). In any case, in the Matthew Project, schools and districts (not individual students) are the object of study (called "the unit of analysis" in the language of researchers).

Tables 1 and 2 provide some information about the diversity that characterizes our four states. Table 1 reports various educational input and process measures, such as percentage os students in smaller schools, Internet connectivity, and several judgments about adequacy and equity (from the prestigious national publication, *Education Week*). Table 2 reports widely accessible state-level aggregate scores on National Assessment of Educational Progress proficiency tests administered in 1996 (eighth-grade mathematics) and in 1998 (eighth-grade reading). The performance of students at this grade level is, in a sense, the acid test for school effectiveness (with achievement the touchstone), because their performance shows the accumulated effects of instruction within the state system but before the attrition of school leaving takes its largest toll in high school. Among the three states that participate in this testing program[2] (and, indeed, nationally) Montana has a substantial history of high achievement, as probed by NAEP.

[2]Ohio no longer participates in the NAEP state-level testing program. However, in the 1992 tests, 59% of Ohio 8th graders scored at or above the basic level in math (Education Week, 1997).

Table 1
Education Input and Process Measures by State[1]

	GA	MT	OH	TX
kids in elementary schools < 350	8	56	24	10
% kids in secondary schools < 900	17	57	49	24
% of schools with class size < 25	64	80	52	90
% schools in need of major repair	26	20	38	27
% **classrooms** with Internet access	35	55	50	42
% **hi pov schools** with net access	87	85	63	79
% all other schools with net access	93	80	81	83
per pupil spending	$4,595	$5,428	$5,438	$4,996
district pp spending disparity[2]	$1,628	$9,171	$5,804	$4,210
education spending ratio[3]	$38	$55	$40	$48
average teacher salary	$35,688	$30,604	$38,833	$35,148
Ed Week Equity Grade	B	B-	C+	D
Ed Week Adequacy Grade	B-	C	B	C+
Ed Week Accountability Grade[4]	A	D	A-	A
Ed Week School Climate Grade[5]	C	B-	D+	C+

Notes. 1. all data from Education Week (1997, 1998)
2. difference between per pupil spending of districts at the 95th and 5th percentiles on spending
3. education spending for every $1,000 of per capita income
4. Education Week "accountability grade" is based on the degree to which states adopt "high standards for all children and assessments aligned with those standards," a perspective that does not enjoy universal support.
5. School climate grades varied from B+ (one state, VT), B (one state, ME) and B- (three states) to D- (three states). 27 states earned grades of C-, C, or C+.

Table 2
National Assessment of Educational Progress Results

	Reading (1998)		Mathematics (1996)	
	% of students testing at or above "basic"	average NAEP score	% of students testing at or above "basic"	average NAEP score
GA	51%	257	51%	263
MT	93%	270	75%	283
OH	---	---	---	---
TX	76%	262	59%	270

Notes: Aggregate NAEP state-level reading scores range from about 230 to 270; aggregate state-level math scores range from about 233 to 284.

Clearly the educational systems differ markedly in these two states--they differ on measures of structure and governance, on measures of resource allocation (process of the education system), and on outcomes in comparison to one another and to the nation as a whole, according, at least, to the National Assessment of Educational Progress, arguably among the most carefully designed of the various accountability assessment schemes. Those who insist on fashioning "nationally representative" pictures of educational processes will apparently miss a great deal of variation associated with between-state differences. We believe that studies that attempt to synthesize a national picture of education miss the point that the US maintains a dramatically decentralized system in which longstanding experimentation and local options have evolved quite different state-based systems of schooling. The results of these experiments should

be of interest to those who make policy, those who teach, and those who vote. The Matthew Project reports bear witness, we believe, to the importance of state-based educational studies.

Synthesis of Procedures for State Studies

Bickel was responsible for analyzing data from Georgia and Texas; Howley was responsible for analyses in Ohio and Montana. Data employed in the analyses varied from state to state, so that results are not strictly comparable. Patterns in the data analyses, however, are sufficiently strong to derive some overall conclusions.

Method. The Matthew project has conducted a series of studies in which equations relate size of schools or districts, average socioeconomic status of those same schools or districts, and the interaction of size and socioeconomic status[3] in order to predict the aggregate student achievement[4] of schools and districts. That is, the performance of schools and districts--not individual students--was what we sought to predict. These equations all look something like this, and are really quite simple:

$$\text{size} + \text{SES} + (\text{size} \times \text{SES}) = \text{achievement}$$

If, in these equations, the interaction term proved statistically significant, we took that fact to mean that the influence of size on achievement varied systematically in tandem with SES. This being the case, we calculated the size of that effect (effect size) using a method pioneered by

---

[3]independent variables

[4]dependent variable

Friedkin and Necochea (1988) and applied subsequently in Howley (1995, 1996).

In some cases, we performed additional analyses in order to help draw out the practical implications of findings. One such analysis that we performed for every state concerns an equity effect of small size of schools and districts. These analyses we performed regardless of whether or not the interaction hypothesis was confirmed. One effect of small size, in view of the interaction hypothesis, is that smaller units mitigate the damaging effects of poverty on achievement. That is, "excellence" is more closely approximated when impoverished communities are served by small schools and districts. In this case, it would seem that the small size helps disrupt the usually strong relationship between SES and achievement. This means, that in impoverished communities, excellence is cultivated via an apparent equity effect (that is, breaking the usual bond between SES and achievement). The equity question is whether or not this phenomenon actually pertains to small schools across the board--*regardless of community SES*. In the Matthew Project we tested this possibility by dividing districts and schools at the median of size (the size that divides the small half from the large half) and computing the correlation (Pearson r) between SES and achievement for each half thus defined.

Other analyses that appear in some of the state-level reports investigate the differences between one group or another on various measures, as necessary to help interpret results of the regression equations. In some cases, as well, we introduce control variables (e.g., pupil-teacher ratio) to see whether such additional variables alter the prediction given by our basic model.

State data sets. In all these state analyses, our "sample" was planned to be all the schools and districts in each state. When an entire group, instead of a sample, is used in analyses, the

calculation of significance levels is sometimes considered superfluous. The reason is that, since sampling error is not at all an issue (all cases are used, so that estimates for a subgroup are not generalized to the entire group), the observed measurements directly and accurately characterize the prevailing relationships. We have, however, retained the use of significance levels, as we believe that nonsignificant ($p>.05$) relationships, almost by definition, are practically insignificant as well. Understand, also, that the process of obtaining, cleaning, merging, and analyzing the data inevitably reduces the actual number of cases available for analysis by a small proportion[5]. The number of schools and districts on which we are able to base our results is somewhat less than the total number of districts in the state, but is much larger than a representative sample would be.

Our dependent variables were in all cases school- and district-level aggregate performance on standardized measures of achievement. In Georgia, test scores came from the Iowa Test of Basic Skills (ITBS); in Montana, test selection is left to the discretion of districts, but scores came principally from just three tests, the ITBS, the Comprehensive Test of Basic Skills (CTBS), and the Stanford Achievement Test (SAT); in Ohio, test scores represented performance on the Ohio Proficiency Tests; and in Texas, the scores represented performance on the Texas Assessment of Academic Success (TAAS). Correlations among state-developed proficiency tests and conventional norm-referenced achievement tests (such as the ITBS or the

---

[5]Missing values on some data and listwise deletion of missing cases (deletion of cases that do not contain values for *all variables* in the analysis) means that the number of cases will also vary from analysis to analysis. We excluded all "special circumstance" districts (n=5), and schools within such districts, from analysis. Special circumstance districts include very small districts on Great Lake islands, for instance; three do not offer high school instruction.

CTBS) are usually moderate rather than strong. Consistency of findings across different sorts of achievement measures, then, would provide unusually strong support for the interaction hypothesis and the hypothesized equity effect.

Our measure of school size was identical in all analyses. We used the total enrollment of a school divided by the number of grade levels contained in the school. This measure has the advantage of controlling for the possibly confounding influence of school grade-span configuration. That is. two schools with 800 students are not truly the same size if one contains 4 grades (say, K-3) and the other contains 9 grades (say, K-8). Our measure of district size was generally the total number of students enrolled in the district, except for some grade-level analyses in Montana, where the state maintains school districts with three different grade-span configurations. For such analyses in Montana, we used the same metric as we did for the school-level analyses.

To measure SES we selected the available measures that correlated most strongly with our achievement measures. For all but Ohio, this was schools' and districts' rates of free-and-reduced-price-meal provision. For Ohio, district-level rates of Aid to Dependent Children were used on this basis. Again, we think that similar results obtained with somewhat different measures from state to state serve to demonstrate the robustness of the hypothesized relationships.

Reporting. Each of the state studies reports its results in details, according to circumstances relevant to the features of the obtained data and the relevant features of schooling. At a minimum, regression equations are reported for both schools and districts, as well as

correlational analyses that measure the comparative strength of association between SES and school and district academic performance for smaller and larger units (i.e., schools and districts). Some of the state reports also provide addi nal analyses prompted by local circumstances. Readers interested in all these details are urged to consult the individual state reports. This national report will merely synthesize implications from patterns of results across states.

Results

Our focus in this report is to point out commonalities across the state reports rather than to recapitulate individual state-level results. Readers are referred to the individual reports, available from the Rural Challenge Policy Program, for details of the findings from each state study.

Excellence effects of size. Based on results from the four state studies, as well as previous research, we would predict that the effect of school size on academic achievement in most states in the nation is contingent on community SES, as Friedkin and Necochea (1988) originally hypothesized. The effect was pronounced in Georgia, Ohio, and Texas. It was not pronounced in Montana, a state that maintains very many small schools. The weakness of the interaction effect in Montana could be the result of this fact.

The inference about Montana is reinforced by a special analysis done for the 132 K-12 schools maintained in Texas. These so-called "unit schools" enroll all the children in grades K-12 in one building and enroll an average of about 270 in 13 grades; otherwise their test scores and SES are comparable to those of all other schools. There is also substantial variability in their

size (standard deviation = 170). Not only is there no interaction effect and no direct effect of size, but the influence of SES on achievement is surprisingly low. One might conclude, as in Montana, that the generally smaller size of these schools eliminates the interaction effect, and, as well, seemingly exhibits an equity effect of small school size.

We discovered a strong interaction effect of district size only in Ohio, though a weak such effect was observed in Montana, for grade 8 students in elementary districts, and for overall district achievement among the few (n=51) K-12 districts in that state. No evidence of any district size effect was found in Georgia, and in Texas, though there was no interaction effect, the direct effect of size was negative for our grade 8 and grade 10 analyses. Apparently, in Texas, increased district size is related to lower test scores for above the elementary level no matter what the district SES.

Equity effects of size. The results for our analyses of the possible equity effect of small size, however, were *stunning*. At *all grade levels*, for *all analyses*, for *different sorts of achievement* and even for *alternative measures of SES*, smaller units exhibit a reduced (often *substantially* reduced) relationship between SES and achievement as compared to larger units. The degree of mitigating effect varied by test, grade level, and state, though it was seldom less than a 20% reduction and seldom more than a 67% reduction. The average benefit seemed to hover between 30% and 50%. Reducing the negative impact of poverty on academic achievement by as much as 30%, however, would be a phenomenal accomplishment.

Other influences on achievement. Since our regression model (see above) was comparatively simple, we also performed analyses that added ethnicity and class size as controls.

The influence of ethnicity on school outcomes, whatever the mechanism, is an important political concern. In our equations for Georgia, Ohio, Texas, and Montana, however, the addition of such controls had very little effect on our reported results. Separate results for schools enrolling high percentages of minority students, however, generally tended to show strong negative *direct effects* of size. Often, in urbanized states like Ohio, impoverished urban African-American communities (but not necessarily impoverished urban or rural white communities) must send their children to large schools. Such results as these suggest that the state-level picture of interaction effects, as one might surmise, is the product of a complex system of allocating benefits and risks differentially among the population according to such characteristics as ethnicity and affluence[6].

We imposed controls for class size as a test of whether or not our school-level effects might not simply be the result of differences in class size. We performed such analyses in Montana and Ohio, without much effect on results. The pertinent regression equations are reported in the Appendices of the reports for those states.

Conclusions

On the basis of replications in 7 states, we hypothesize that school performance in the various is *widely characterized* by an interaction effect of size, such that smaller schools benefit impoverished students and larger schools benefit more affluent students. This effect probably

---

[6]With benefits generally accruing principally to those whom one might suspect, i.e., not principally the poor and not principally ethnic minorities like African Americans.

cannot be found in states that operate mostly small schools (e.g., Montana), or mostly large schools.

We also conclude, with substantial confidence, that an equity effect of small size characterizes academic results in districts as well as schools. The smaller the unit, the weaker the bond between unit performance and unit SES. This equity effect would be further hypothesized to pertain to all common SES and achievement proxies. Evidence for this effect, we repeat, is unusually strong in the Matthew Project studies (see also, Howley, 1995, 1996).

We also conclude that these results are not the artifact of excluding other influential variables such as ethnicity or class size, nor are they due to anomalies of data (such as skewness or use of inappropriate measures.

The Matthew Project studies provide strong evidence that a one-best, everywhere "optimal," school size is a figment. The appropriate size for a school, when the aim is to maximize aggregate student achievement, depends on community circumstance, operationalized here as aggregate SES. For very impoverished communities, large schools would, on the basis of the reported findings, be expected to produce educational impoverishment, not educational enrichment.

We also infer from our equity findings the notion that schools can be so large as not to serve anyone very well. This is to claim that it might be wise to establish some upper limit of school size, even for schools serving very affluent communities. Recent incidents of carnage in ostensibly peaceable suburban schools may speak to the need for such wisdom.

Discussion and Recommendations

A key question for states like Alaska, California, Georgia, Ohio, Texas, West Virginia (and to a much more limited degree even Montana) is the source of the interaction effect identified by study results. The evident result could be produced (as in West Virginia) mostly by the presence of many small schools serving impoverished communities, or (as in California) by many large schools serving impoverished communities.

Or it could be a combination of conditions, as seems the case in diverse, urbanized Ohio. In some cases this diversity will probably have supported equity and adequacy of outcomes, but in other cases it will probably have undermined such outcomes. Ohio resembles California in that the majority of poor African-American students are served by large urban schools, but many small rural schools serve (largely White) impoverished communities in southeast Ohio, while still others serve communities with varying degrees of affluence.

Whatever the *interesting details*, however, they cannot generally be of great help to those who fashion policy and to those who are concerned to maximize the intellectual potential of all citizens. In fact, neither researchers nor the public should look to research to *determine* answers to the important and difficult questions, but only to *inform* them.

The minute details may be interesting, but are not critical, at least in this instance, to public policy. The overall findings are more relevant to policy and practice:

(a) ***one size cannot fit all*** and
(b) ***smaller units mitigate the negative effects of poverty.***

The forest, in this case, is more important than the trees, at least with reforestation in view. The Matthew Project results, however, point to contradiction. so *enlightened* public debate (one based on the preceding overall findings) will help clarify matters.

Citizens and policy makers in the states might consider four questions related to the issue of district and school size. Though not answered very easily, the important questions are stated quite simply:

1. Should states set an absolute upper limit on the sizes of districts and schools? If the answer is "yes," then what should these caps be for elementary, middle, and high schools and for districts?

2. In states where an interaction effect exists, which districts and schools violate these limits? What should be done in these cases?

3. In states where an interaction effect exists, how should such upper limits vary by school and district SES? Should *any* lower limits prevail? What would they be?

4. What policies might successfully promote the re-scaling ("restructuring") of state systems of schooling as needed the better to serve all the communities whose children attend them?

One of us has, with respect to the first question, suggested and illustrated a logic for establishing such upper limits as they effect *school* (not district) size (Howley, 1997). When these suggested limits were, for instance applied to Ohio schools, approximately 30% of high schools, 40% of middle schools, and 50% of elementary schools were shown to exceed the limits

derived from this logic.[7] For Montana, the comparable results were 8%, 4%, and 7%.

We cannot, of course, maintain that the suggested limits are the best, the most sensible, or the most logical. They are eminently defensible, however (see Raywid, 1999)[8]. This example, however, shows that such considerations can establish benchmarks capable of revealing striking differences that should provoke some debate about states' policies on school and district size, in light of results such as those reported in the paper.

The fourth question, above, is perhaps the key policy question. Because it revolves around the unfamiliar issue of scale, however, it is more complex than it might at first seem.

The concept of scale is most fully described in chaos theory, strictly and not loosely defined (e.g., Gleick, 1987). In chaos theory, the evident "chaos" of appearance (as in weather and climate patterns, population growth, plant growth, hydraulic flows, and so on) is actually ordered by the simple repetition of characteristic structures at different levels of detail. Computers have allowed us to appreciate this previously obscure reality. The now classic general example of this iterative patterning of the structure of chaos is the Mandelbrot set, which repeats itself endlessly the finer the detail at which it is examined. This is the phenomenon of "scale" that might apply to state systems of schooling.

---

[7]The logic derives from the fact that elite private schools enroll about 1,000 students in grades 9-12, taken as the upper limit based on the notion that a school this large is most suitable for the most affluent community. Upper limits for common grade configurations in Ohio based on Howley (1997) are: 9-12 (1,000); 6-8 (600); and K-6 (400). The most common middle-level configuration in Montana is 7-8 rather than 6-8, so 400 students would be the hypothetical upper limit. I have not suggested upper limits for district size, however.

[8]Raywid reports that upper limits recommended in the recent professional range from 500-900, which would make our recommended upper limit quite conservative.

For instance, the scale of operations in Montana, as compared to Ohio, would (hypothetically at any rate) seem more finely adjusted and better balanced: in effect, more intimate and humane overall. The state system is much smaller as a whole--serving about 1/10 the number of children as compared to Ohio. The districts and schools are smaller, as noted previously, and class sizes are smaller than the national average as well. In fact, the very smallest Montana classrooms retain the multi-age grouping of single-teacher schools.[9]

Chaos theory would suggest, however, that what one might refer to as the "more intimate" scale of public schooling in Montana extends to a still more finely grained level *within individual students and teachers*. This possibility is highly speculative, of course, and strange, so please bear with the discussion. The hypothetical reality of "small scaling" within individuals might be understood as a possible matter of *attention* (i.e., to people, to facts, to ideas, to dispositions, and to relationships) characteristic of more "intimate" or "better balanced" or "more finely adjusted" contexts.

Unfortunately, people within such an "intimate" system are no better placed to recognize their circumstances than a tree within a forest.[10] This assertion, which, incidentally is a common one throughout the history of science, would explain why we have not previously

---

[9]That is, with one or two students per grade, such districts, whether elementary or high school or K-12, surely do not maintain separate classrooms or teachers for each grade, as the norms of contemporary professional practice (age-grade-placement) would otherwise require.

[10]Rural people, do, however, express a sensibility of this sort--which might otherwise be called "sense of place," "connection to the land," or when lost, yearning. The eminent sociologist Christopher Lasch reminds us that the word "nostalgia" emerged from the experience of displaced rural populations in the nineteenth century, who, much like the urban Appalachians of today, keenly suffer the loss of a sense of place.

noticed the implications of scale effects in systems of schooling.[11]

Though this speculation is risky, we offer it primarily to suggest the ambiguity and complexity of answers that might possibly given to question four. Answers to the first three questions would be simple by comparison to any more or less fully given answer to question four. The fourth has a "revolutionary" teleology, with implications for turning state systems of schooling upside down or inside out, in the very unlikely event the teleology were actually pursued even a bit.

Limitations

Precise comparison of results across states is simply not possible in this series of studies due to lack of comparable data for all schools and districts in our four states. Nonetheless, given the prevalence of the interaction effect for school size and the highly consistent equity effects of school and district size, this shortcoming also constitutes a strength of sorts. When similar findings are discovered under somewhat different methodological as well *and* real circumstances, we can have greater confidence in them.

The shortcoming, of course, is that we cannot precisely specify the magnitude of the effects, generally, nor make very precise comparisons of differences in the magnitude of effects between states. This problem is related to a very common problem in social science research; that is, the need to operationalize variables at all. Constructs such as achievement, school and district size, and socioeconomic status can *never* be precisely measured at all, but only

[11]Though ultimately accessible, reality is seen to lurk well beneath the surface of appearances.

approximated by better or worse proxies. The proxies used in these studies are defensible operationalizations of the constructs, but not perfect ones. Better proxies, especially for SES, might well yield effects still stronger than those reported here.

Finally, the Matthew Project studies *do not* address the distribution of achievement among individual students within schools. Comparatively good school- or district-level aggregate scores can be had through various combinations of improved achievement among different groups of individuals in the same school or district. Improvement might come at the hands of high- or middle- or low-achieving students principally, or from some combination of these. *For most grade-level tests (as these are), however, the likeliest way that aggregate scores improve is with improved scores among otherwise low-performing students.* Test construction is the reason. With grade-level tests, high-performing students will *already* be performing closer to the rather low ceilings of these test than other students. This simple fact means that improvement among high-scoring students is more chancy--that is, less likely--with than among low-scoring students simply because the opportunities for such improvement dwindle as scores approaches the test ceiling.[12] Thus, we would reasonably hypothesize that smaller schools enable the "bottom half" of the achievement distribution to make a closer approximation to "normal" achievement levels.

Curiously, this observation might, as well, suggest that the observed interaction effect *underestimates* the beneficial effects of larger size on more affluent students. If such students were administered wide-range (rather than grade-level) achievement tests, they might

[12]That is, fewer items sample the higher reaches of performance, so that scores become more unreliable, and the error bands around individual scores widen.

demonstrate more reliably a level of achievement beyond that which any grade-level test can reliably measure. These possibility should be pursued in future work.

Theoretical Perpsective: An Addendum

In theoretical terms, previous studies of school size have taken a functionalist approach, seeking and finding *equilibrium* in the system of schooling. In the present work, however, the issue of school size is seen as an issue of "contradiction," a concept from the structuralist perspective[13].

The school-closure and district reorganization battles of the past century would seem to warrant a structuralist perspective. A structural view (i.e., rather than only a "structuralist" view) of the issue of size is also consonant with the literature on private enterprise, however. As in much of the work about firms, in the present work, the size of districts and schools is understood as a durable condition (i.e., a structure) and *not* as a container of effective processes that might be extirpated and transplanted elsewhere (e.g., in larger schools) or as necessarily harmful ("small is good") or necessarily beneficial ("bigger is better") condition. *The line of inquiry carried on by the Matthew Project continues to entertain the unusual possibility that smaller schools and*

[13]"Contradiction" is a structuralist notion particularly apparent in various qualitative views of social organization. Contradictions manifest themselves as logical incompatibilities (e.g., small schools serving poor students best and large schools serving large students best even as all students must attend a single school), but, in a structuralist epistemology, are taken to reflect the dynamics of social structure from which social change arises. Thus, a century of school closures might hypothetically suggest a deployment that has tended to serve the affluent better than it has served the poor. Closures, however, nearly always produce conflict and contention as the less powerful object to their treatment. On this view, one might say that the social contradiction has developed to the point that it can actually be gauged with a functionalist research technique (i.e., regression analysis with a salient interaction term). Of course, this (methodological) contradiction is (itself) related to many others in the social structure.

larger schools might be beneficial in some circumstances but not in others. As always, when power is brought to bear, the question is *who benefits* from actions taken by powerful individuals (e.g., politicians) and institutions (e.g., SEAs) and *who suffers.*

Today, especially, it would seem that state school leaders might anticipate an improved efficiency of reform in a more consolidated state system. Simply communicating with over 500 districts about the minutiae of legislation and new initatives is an understandably daunting chore.[14] In the age of systemic reform, the longing among state officials for fewer and fewer districts must indeed be strong. Of course, educational equity (measured as parity in district-level school funding) is improved when there are fewer districts. States with fewer districts (as in the South) therefore appear to be more equitable (Education Week, 1997). Whether or not such equity is merely or mostly an artifact of organizational structure has not been carefully questioned. Such equity, in any case, hardly assures either adequacy or equity of outcomes.

[14]In the rural South, enduring post-Bellum (i.e., Civil War) rural poverty has ensured careful attention to *financial* efficiency, a commitment that one might argue has ensured more prevalent consolidations and closures than has been possible elsewhere. However, in places settled by independent "yeoman" farmers, consolidation has been far less successful. Illinois, Missouri, Ohio, Nebraska, and Pennsylvania all maintain more than 500 LEAs to this day (as do Montana, Ohio, and Texas).

References

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.

Education Week. (1997). *Quality counts: A report card on the condition of public education in the 50 states* (supplement to *Education Week*, January 22, 1997, Vol. XIV). Washington, DC: Author and the Pew Charitable Trusts.

Education Week. (1998). *Technology counts: Putting school technology to the* test (*Education Week, 18*(5), special issue). Washington, DC: Author and the Milken Exchange on Education Technology.

Friedkin, N., & Necochea, J. (1988). School system size and performance: A contingency perspective. *Educational Evaluation and Policy Analysis, 10*(3), 237-249.

Gleick, J. (1987). *Chaos: Making a new science.* New York: Penguin Books.

Goodlad, J. (1984). *A place called school: Prospects for the future.* New York: McGraw-Hill.

Howley, C. (1995). The Matthew principle: A West Virginia replication? *Education Policy Analysis Archives* [On-line serial] *3*(18). (Available World Wide Web: http://olam.ed.asu.edu/epaa/v3n18.html)

Howley, C. (1996). Compounding disadvantage: The effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education, 12*(1), 25-32.

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. (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.

Huang, G., & Howley, C. (1993). Mitigating disadvantage: Effects of small-scale schooling on student achievement in Alaska. *Journal of Research in Rural Education*. 0(3), 137-149.

Kahlenberg, R. (1999, March 31). Commentary: Economic desegregation. *Education Week* p. 52-30.

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. 425 049)

The Matthew Project

The Matthew Project, with funding from the Rural Challenge Policy Program (now known as the Rural School and Community Trust Policy Program), investigated the possible academic excellence and equity effects of school and district size in Georgia, Montana, Ohio, and Texas. The project title refers to a parable about stewardship in the gospel according to Matthew (13:12): "For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath." Building on previous research efforts in Alaska, California, and West Virginia, the Matthew Project was particularly concerned to investigate the possible contributions of smaller school size to academic success in impoverished communities.