

Is Politics the Problem and Markets the Answer? An Essay Review of *Politics, Markets, and America's Schools*

ANTHONY S. BRYK* and VALERIE E. LEE†

* Department of Education, University of Chicago, 5835 S. Kimbark Avenue, Chicago, IL 60637; and
† School of Education, University of Michigan, Ann Arbor, MI 48109, U.S.A.

Abstract — *Politics, Markets, and America's Schools* is an ambitious book that draws eclectically on concepts from political science, the sociology of organizations and educational learning theory. Chubb and Moe employ an extensive array of data seeking to link the mechanisms that control school operations to student achievement. They conclude that a total restructuring of the governance system of American education — from democratic to market control — is necessary. Our review scrutinizes a set of critical decisions made by Chubb and Moe in defining their key concepts and in the analytical models employed in this research. We argue that many of these decisions are not justified on either theoretical or methodological grounds. Moreover, the cumulative effect of the decision tends to tilt the empirical evidence toward supporting the authors' *a priori* beliefs. As a result, we conclude that *Politics, Markets, and America's Schools* is best viewed as a policy argument, where extensive, but not always solid empirical evidence has been artfully employed to advance the authors' preconceived notions about American schooling.

1. INTRODUCTION

A QUICK SCAN of *Politics, Markets, and America's Schools* suggests an impressive book. Over the course of some 300 pages, John Chubb and Terry Moe craft an empirically grounded argument demonstrating that how a school is governed influences its internal operations which in turn affects student learning. Opening with a discussion of popular ideas for understanding the current malaise of American education, the authors proceed in Chapter 2 to lay out an alternative conceptualization which is based on an institutional perspective on schools. They premise their theorizing on past research which has demonstrated that bad schools have weak organizational properties. They reason that, "If one really believes that schools are open systems (a common claim), bad organizational properties must be understood as symptoms rather than causes. The fundamental causes are probably in the environment, and it is there that theory and research on school effectiveness ought to focus" (p.

19). Ultimately, this provides the rationale for scrutinizing school governance.

The basic argument — that the root of America's educational problems are in the democratic control of schooling — is supported by an extensive array of empirical evidence in Chapters 3, 4 and 5. Chapter 3 documents that the organization of high and low performing schools is quite different. Effective schools having clearer goals, better leadership, more emphasis on academic coursework and a higher level of teacher collegiality, influence and status. Chapter 4 demonstrates that the effective organizational properties identified in the previous chapter are linked to students' academic performance. The argument culminates in Chapter 5, where Chubb and Moe muster evidence that the key predictor of effective organization is school autonomy, and the critical determinant of school autonomy is whether the school is private (i.e. subject to market control) or public (i.e. democratically controlled). Ergo markets should replace democratic control of schools.

The goals of *Politics, Markets, and America's Schools* are extensive. From a theoretical perspective, the book draws eclectically on concepts from political science, the sociology of organizations and educational learning theory in seeking to link school control to student achievement. Empirically, these constructs are operationalized with data from a large national survey of high schools and students, *High School and Beyond* (HS&B), and a special *Administrator and Teacher Supplement* (ATS). The construction of measures and statistical analyses were among the more extensive encountered in educational research. From a policy perspective, the book's aims are also sweeping. The authors argue that nothing less than a total restructuring of the American education system — from democratic control of public education to a market system — will suffice to redress current problems.

The authors should be applauded at the very least for their ambition. By virtue of the book's intent, and the media coverage that has surrounded this intent, the authors have certainly captured public attention. We believe, however, that the book falls considerably short of its stated aims. From a purely scientific perspective, we doubt whether, having gained the public's attention, Chubb and Moe really have much to say that is new. In our view, *Politics, Markets, and America's Schools* is best viewed as a policy argument, where extensive, but not always solid, empirical evidence has been artfully employed, to advance the authors' preconceived notions about American schooling.

The breadth of conceptualizing and wide scope of the empirical detail covered in this book pose a serious challenge to the critic. At best, a review such as this can select a few (hopefully important) points from this much larger tapestry for scrutiny and comment. We focus our remarks on a combination of technical and conceptual decisions about the key concepts in the study — student performance, school organization, and external control — that, we believe, materially affect the authors' conclusions.¹

2. CONCEPTUAL AND TECHNICAL ISSUES RAISED IN THE CHUBB AND MOE ANALYSES

In the course of such a complex study as this, researchers routinely make hundreds, perhaps thousands, of technical decisions about matters such as variable construction, sample definition, analytic

approach and selection of results during report writing. If this process is guided by appropriate theory and constrained by established methodological practices, we can assume that while some decisions may favor the theory under study, others are likely to work against it. Overall, we expect a fair test of the basic propositions under investigation.

On the other hand, if researchers are inclined to craft and defend a particular argument, these numerous decisions can often be shaded, sometimes in very subtle ways, to support their case. Because of the unusual sequence of events surrounding the publication of this study (see conclusion) and the way evidence is presented (or not presented) in the book itself, we are inclined to examine their methods a bit more closely.

3. SELECTION OF THE KEY OUTCOME: A COMPOSITE MEASURE OF STUDENT PERFORMANCE

Distinguishing between student achievement and aptitude has been a longstanding concern among educational psychologists. Chubb and Moe's own resolution of these nettlesome issues is to employ a composite of five sophomore year test scores collected in HS&B which they argue, "offers a suitable measure of student ability" (p. 116). The *difference*, however, on the same composite measure between sophomore and senior year, they declare is a highly reliable index of "total student achievement" (p. 71). Although Chubb and Moe often use the terms aptitude and achievement interchangeably, they ignore the fact that these are fundamentally different concepts. It is generally accepted that schools should be held accountable for the latter but not the former.

We have argued elsewhere (Lee and Bryk, 1988) that the only valid measure in HS&B for studies of school effects on student achievement is the mathematics test. Mathematics is a core subject in high schools, unlike reading and vocabulary (two other HS&B tests) which generally are not directly taught. The content of the mathematics test also bears at least some relation to typical high school instructional objectives in grades 10–12, not generally true of the rest of the HS&B test battery. While other school effects researchers have taken a more liberal view on this point, virtually all eschew using an HS&B composite measure in favor of the

subject-by-subject analyses (Coleman *et al.*, 1982; Gamoran, 1987; Hoffer *et al.*, 1985).

At base here is a simple but important idea. Student learning is not like the Gross National Product. There is no theoretical justification for a concept of "total student achievement" (p. 71) as rationalized by Chubb and Moe. Students learn individual subjects especially when they are taught those subjects.² In our view, the composite measure is best viewed as a general ability indicator analogous to an SAT score. While it surely contains some "school effect", it is not a particularly sharp instrument for such investigations. Instead, the choice of the composite measure serves to tilt explanatory power in this study toward global school characteristics like autonomy and overall organization and away from specific instructional practice and aspects of teacher quality.

In general, the quality of teachers and instructional practices are quite variable within schools. As a result, these concerns typically require a specific subject matter and classroom focus to investigate them properly. HS&B data are inappropriate by design for such studies because HS&B reports no classroom level data. At best, it provides aggregate information at the school and teacher level which is of little value in studying the effects of classroom constructs, since enormous variation exists among teachers within schools and even for the same teachers among their various classroom assignments (Raudenbush *et al.*, in press). Thus, the conclusions offered by Chubb and Moe that "the classrooms of successful and unsuccessful schools often differ very little. There is a remarkable sameness to the classrooms of America's schools" (p. 96) are simply unwarranted.

4. AN ODD METRIC: A LOG TRANSFORM OF AN IRT SCORE

Typically, HS&B school effects studies have used raw test scores (i.e. the number of items correct) as the outcome variable metric. Through the use of a rescaling technique, called item response theory or IRT (Lord, 1980; Hambleton, 1989), the differential difficulty, discriminating power and likelihood of correctly guessing various items are taken into account to form an *interval scale* measure. Basically, the IRT score is a non-linear transformation of the number of items correct, which has the effect of expanding the scale at its extremes. For example,

one or two more items correct implies a much bigger change in IRT scores at the top of the scale than in the middle. A consequence of this rescaling is that the estimated gains between sophomore and senior year for initially high achieving students will appear larger in an IRT metric than in the raw metric. In the Chubb and Moe analyses, this choice of metric favors private and suburban schools, whose students typically scored considerably higher on the sophomore tests.

Some controversy still surrounds the theory and use of IRT, but it has become increasingly commonplace in major testing programs such as the National Assessment for Educational Progress. Although this metric is likely to work to the advantage of Chubb and Moe's argument, its use is technically defensible. However, Chubb and Moe were apparently not satisfied with this metric and proceeded to transform the measures a second time. The acknowledged purpose for this second transformation was to adjust the estimated student "improvements for the level at which improvement begins" (p. 76), thereby expanding even further the size of the estimated gains in the initially high achieving schools as compared to more average schools.

As noted above, an explicit purpose of IRT is to create an interval scale measure. This means that any difference among two IRT scores has the same meaning regardless of where they are located on the scale. Thus, a simple difference in IRT scores is a proper gain score metric. No statistical justification exists for further rescaling the IRT scores to adjust for "the level at which improvement begins". This is, however, another technical decision that tilts in favor of Chubb and Moe's argument.

In their defense, the authors might claim that the analyses reported in Chapter 4 takes into account this over-adjustment by introducing the 10th grade composite scores as a covariate to "adjust for initial ability". The authors appear unaware that a gain score analysis controlling for the pre-measure is formally equivalent to simple analysis of covariance (ANCOVA) of the 12th grade scores controlling for the 10th grade measures.³ Thus, the end result of these various data manipulations is an ANCOVA but now where the outcome is in a transformed metric that favors advantaged schools.

5. THE HIGH PERFORMING SCHOOLS: WHO ARE THEY?

The ATS involved a sub-sample of about 400

schools from the larger HS&B survey. The various sampling procedures employed here result in a complicated set of design weights whose appropriate use is not always obvious, particularly in complex analyses as presented in this book. These problems are further amplified by some of the specific analytic strategies employed by Chubb and Moe. For example, the analyses in Chapter 3 use a post-hoc stratification into high- and low-performing schools, defined respectively as the top and bottom quartile on composite student achievement in the ATS sample. At best, the ATS school weights apply to the whole ATS sample.⁴ When a post-hoc stratification is imposed, as in this case, there is no assurance that the use of weights within quartiles provides representative data of high- and low-performing schools in the U.S.A., as claimed by Chubb and Moe. At a minimum, we should approach with caution descriptive comparisons based on the two extreme quartiles.

Curiously, while these comparisons of high- and low-performing schools are an essential part of their argument about the probable causes of student achievement, some very basic facts about these groups are not reported in Chapter 3. Where are these schools located and how are they governed? In fact, it is not until the very end of their empirical work, late in Chapter 5, that private control and urban location are mentioned. From the results presented in Chapter 5, we infer that the high performing group consists of a melange of relatively large suburban schools and small private schools. (The latter accounts for the odd bimodal distribution reported in Table 4.2 for school size in the high performing group.) This subset of schools includes some very different organizational forms. At one extreme are the sophisticated suburban "shopping mall" high schools (Powell *et al.*, 1985) with their large and diverse curricular offerings (which offer something here for everybody). At the other end, are small private schools addressing a unitary purpose (see for example, the core academic curricular emphases of Catholic high schools described by Bryk and Lee, 1988, 1989; Bryk and Lee, in press; Hill *et al.*, 1990). Suburban high schools have among the highest levels of fiscal resources in the country, and some of the religious private schools among the fewest.

These observations are not new. The distinctiveness of the Catholic sector and other private sectors was well documented in the very first study using

HS&B (Coleman *et al.*, 1982). Evidence has accumulated that the structure of relations within schools varies as a function of school governance, school social class and other characteristics (see for example, Coleman *et al.*, 1982; Lee and Bryk, 1989). In all likelihood, the mechanisms that produce effectiveness in affluent suburban public schools are quite different from those at work in smaller religious private schools. Thus, it is just not sensible to lump these schools together. The homogenization that results from this aggregation again works to mask the influence of differential internal organizational factors in favor of more macro external factors — another critical analytic decision that favors Chubb and Moe's argument.

6. ATHEORETICAL CONCEPTUALIZATION AND MEASUREMENT OF SCHOOL ORGANIZATION

The quality of school organization plays the central role in Chubb and Moe's argument. Theorists such as Durkheim (1933, 1956, 1961), Dewey (1900, 1966, 1981) and Waller (1932) have brought attention to how roles, relationships and rules work in schools. Others such as Weber (1947), Parsons (1959) and Bidwell (1965) have written extensively about more macro aspects of school structures and norms. Taken together, this work affords a rich theoretical base for investigating how school organization might affect student learning. Unfortunately, little of this theory is employed by Chubb and Moe in their measure constructed for the school organization concept which they describe as a key measure in their work (p. 120).

Instead, Chubb and Moe pursue a more empirical approach. They consider a range of organizational characteristics that might be related to student performance including school goals, principal leadership, personnel quality and selected instructional practices. The features identified as differing between low-performance and high-performance schools were combined to create two composite indices of school organization: a comprehensive index with 10 components, several of which are themselves composites; and a condensed index with seven components. The rationale for creating the condensed index from the comprehensive index was also largely empirical. The condensed measure worked better than the comprehensive measure as a predictor of student achievement.

Of particular significance, the atheoretical approach employed by Chubb and Moe leads them to include in both the comprehensive and condensed indices two components which are not organizational *properties* but rather organizational *consequences*. Specifically, two of the three components used to measure teacher professionalism — teacher efficacy (itself a composite of two items) and teacher absenteeism (a composite of four items) — are generally viewed as teacher outcomes (see for example, Bryk and Driscoll, 1988; Newman *et al.*, 1989; Lee *et al.*, 1991; Raudenbush *et al.*, 1992). The inclusion of these outcomes within the measure of school organization is not inconsequential. Previous research (Ashton and Webb, 1986; Rosenholtz, 1989; Firestone and Rosenblum, 1989) has demonstrated that these teacher outcomes are interconnected with student outcomes including academic achievement. While teacher efficacy and absenteeism comprise important dimensions of both the comprehensive and condensed versions of the school organization measure, the problem is especially serious in the condensed measure, since there are fewer other components. It seems likely that the stronger effects reported for the condensed than for the comprehensive measure (e.g. compare results in Tables 4–10 and 4–11) on gains in achievement are explainable by the larger influence of the two inappropriate components in the condensed version of the variable. In essence, Chubb and Moe have included school outcomes in defining their central predictor of school outcomes.

Moreover, in moving from the comprehensive to the condensed measure, an important dimension of school organization — teacher collegiality — was excluded. The authors justify this decision on the grounds that this was ‘clearly the most subjective indicator of organization in the original measure’ (p. 136). Other recent research using the same HS&B data has reported significant relationships with teacher collegiality and concluded that it is an important dimension of school organization (Bryk and Driscoll, 1988; Bryk and Lee, in press; Bryk *et al.*, 1990; Newmann *et al.*, 1989; Lee *et al.*, 1991). In sum, Chubb and Moe’s condensed measure of school organization includes two inappropriate components (teacher efficacy and teacher absenteeism) and excludes one (teacher collegiality) that should be there. As such, the index suffers from a lack of construct validity — it does not measure

what it purports to measure (Cook and Campbell, 1979).

7. UNDER-REPRESENTING THE EFFECTS OF ACADEMIC ORGANIZATION

The proportion of students in the academic curriculum track is the most important component in the two organizational indices mentioned above. It has the strongest individual relationship to student achievement. This component is described by Chubb and Moe as “gauging the educational practices of the school — in particular, how students are assigned to classes and what takes place in those classes” (p. 122). In analyses which investigate the influence of the school organization index on student achievement, the estimated effects are dramatically reduced once the proportion of students in the academic track is removed from the overall measure (compare Table 4–1 vs Table 4–10 for the entire student sample; Tables D–1 vs D–4 for public school students only). The magnitude of the coefficient associated with the proportion of students in the academic track in Table 4–10 and D–4 actually exceeds that of the residual school organization measure with tracking removed.⁵

Moreover, we contend that the true impact of a school’s academic organization on academic achievement is under-represented by Chubb and Moe. Recent field research has reported the decreased salience of curriculum tracking in high schools, where a large array of courses of varying content and difficulty now represents current practice (see, for example, Powell *et al.*, 1985). Studies investigating school effects on student achievement have been quite consistent in demonstrating that the most important single factor is the courses students have taken *in the particular curricular area being tested*. Similarly, our extensive work comparing the organizations of Catholic and public high schools (Bryk and Lee, in press; Lee and Bryk, 1988, 1989) has demonstrated the primary role of academic organization in explaining the advantageous outcomes (higher achievement and a more equitable distribution of achievement) evidenced by students in Catholic high schools. The schools are characterized by a constrained set of largely academic offerings which is taken by all students, regardless of academic or social backgrounds. To be sure, this advantageous academic organization means that a

larger proportion of students is in the academic track. However, the largest advantages for Catholic school students compared to their public school counterparts, in terms of both academic course-taking and achievement, actually accrue for students in the non-academic tracks. This confirms that it is the courses students take which make the difference, not just their track placement.

In sum, Chubb and Moe failed to adequately measure the single most important component of school organization — the actual academic experiences of students — even though these data exist in HS&B and have been extensively analyzed by other researchers. This analytic decision, not to use specific course taking data in favor of a more global measure, is the organizational parallel to the decision discussed earlier to use a composite measure of student achievement. In both cases, by relying on global measures, the explanatory power of their model tilts towards more general factors external to the school — still more critical decisions favoring the Chubb and Moe argument.

8. SINGLE-LEVEL METHODS AND A MULTI-LEVEL CONCEPTUALIZATION

As noted earlier, the central proposition under research by Chubb and Moe, that properties of school organization and their external environment affect the learning that occurs within schools, is multi-level. That is, it involves concepts defined at both the school and student levels and seeks to trace the relations across the two levels. The problems of multi-level analysis — misestimated standard errors, confounding of within- and between-unit relations, and failure to model heterogeneity of within-unit relations — have been well-documented (for a thorough review see Burstein, 1980). Fortunately, more appropriate statistical models for this phenomenon have emerged in the last 10 years, called multilevel analysis (Mason *et al.*, 1983; Goldstein, 1987) and elsewhere termed hierarchical linear models (Raudenbush and Bryk, 1986; Bryk and Raudenbush, 1992). These new techniques resolve longstanding methodological difficulties, and their use is rapidly becoming common place (see, for example, the edited volume of educational applications by Raudenbush and Willms, 1991). Chubb and Moe, however, have relied exclusively on the more questionable single-level analyses.

Specifically, their key results derive from student-

level analyses of academic performance where a number of student characteristics are entered as controls. It has been previously well-documented that these relations are not homogenous across schools (see for example, Coleman *et al.*, 1982; Lee and Bryk, 1989). Ignoring this heterogeneity, as Chubb and Moe do, introduces a potential bias in interpreting school effects. That is, when schools affect students differentially and the structure of these “within-school effects” varies across schools, then the “average effects” estimated by the Chubb and Moe model can be misleading. From the perspective of analysis of variance models, this is tantamount to interpreting main effects and ignoring the presence of significant interactions. At a minimum, Chubb and Moe should have explicitly recognized the previously demonstrated sector differences (public vs Catholic vs other private schools) in their single-level models and explored possible interaction effects as a function of sector.

This problem is further exacerbated by the weighting scheme deployed by Chubb and Moe. Since their key analyses are student-level, they use the regular HS&B student-level design weights. However, these weights do not apply to the subsample of students in the ATS schools, only to the full HS&B design (the ATS is not a simple random sample of HS&B). Further, the statistical evidence that bears most directly on the authors’ central claims involves the relation of school characteristics to adjusted mean achievement differences among schools. These are school-level relations and the proper weights here are the ATS school weights, not the HS&B student weights.

Short of reanalyzing the actual variables constructed by Chubb and Moe, it is impossible to determine whether the choice of weights works for or against their argument. It is clear, however, that the choice can have significant consequences. To illustrate this, we present results from a small analysis problem, using a subset of HS&B data that simulates the basic features of the Chubb and Moe research — a combination of student and school-level measures predicting 12th grade mathematics achievement.⁶ (See Table 1 for a description of the sample and variables). We analyzed these data at the student-level using both the HS&B student and school weights. The student-level effects reported in Table 2 are quite similar for the different weighting conditions. The school-level effects, which parallel those of prime interest to Chubb and Moe, however,

Table 1. Description of sample and measures for weighting illustration

Sample — drawn from the base-year and first follow-up HS&B

Schools

- All Catholic schools (*N* = 83)
- Random sample of public schools (*N* = 95)

Students

- All HS&B students within these schools who were sophomores in 1980 and seniors in 1982
- 2050 Catholic school students
- 1883 public school students

Oversampling (from HS&B)

- 33 oversampled Catholic schools (containing high proportions of either blacks or Cuban Hispanics)
- 12 high-Hispanic enrollment public schools
- 45 of 178 schools are from oversampled strata

Measures

Outcome

- Senior-year mathematics achievement (FYMTHIFS)

Predictors

- Social class (BYSES)
- Minority status (from sophomore-year measures of race and ethnicity, 1 = black or Hispanic, 0 = everyone else).
- Sector (1 = Catholic, 0 = public)
- Average school SES (aggregated from BYSES)
- % Minority in school (aggregated from minority status)

Interaction terms

- SES × AVSES (product term)
- MINORITY × %MINORITY (product term, with the minority dummy variable effects coded, 1 = minority, -1 = non-minority)

HS&B, High school and beyond.

Table 2. Regression coefficients for school effects model on senior-year mathematics achievement, under different weighting conditions*

	Without weighting [1]	Weighting conditions Raw student weights [2]	Raw school weights [3]
Student level variables			
Social class	1.06 (0.10)	1.18	1.22
Minority class	-3.67 (0.50)	-4.08	-3.93
School-level variables			
Catholic sector	1.64 (0.25)	1.67	1.53
Average school SES	1.25 (0.19)	0.97	0.79
% Minority in school	-2.60 (0.76)	-3.27	-4.71
Contextual variables			
SES × average SES	0.078 (0.09)	0.03	-0.10
Minority × % minority in school	2.67 (1.01)	2.89	3.59
% Variance explained	0.19	0.19	0.19

* The regression coefficients presented here are unstandardized. The dependent variable is a mathematics achievement test, containing 38 items. The overall mean (unweighted) of the test for this sample is 20.9, and the standard deviation is 7.7.

are different. The Catholic sector effect with school-level weighting is 10% smaller, the effect of school social class is almost 20% smaller, and the effect of percentage minority enrollment increased by about

45%! Differences of this magnitude are sufficient to alter conclusions about the significance of school effect such as the results Chubb and Moe report for school organization. As this evidence about school

organization is critical to their argument, an improper choice of weights could result in the whole argument collapsing.

9. ANOMALOUS SCHOOL SIZE RESULTS

Other school organization studies, using these same data, have found that school size is a powerful predictor of the nature of the social organization found in American high schools (Bryk and Driscoll, 1988; Driscoll, 1989; Newmann *et al.*, 1989). Similar results have also been reported in other data. For example, in a recent study of Chicago school reform (Easton *et al.*, 1991), teachers in small schools were much more likely to characterize their work conditions in positive terms. From a theoretical perspective, the influence of organizational size on bureaucratization has been forcefully argued by Weber and others. Thus, school size represents an important rival hypothesis to governance control for understanding organizational effectiveness.

In Table 5.1, Chubb and Moe describe schools with effective organizations as being somewhat smaller, although they argue that the size of these differences is not meaningful. The basis for this conclusion is again a comparison of characteristics of schools in two extreme quartiles, in this case on a composite measure of school organization. The

weighting issue noted above, along with the failure to directly represent school sector (public vs private) in these analyses, may account for these anomalous results.

In fact, the results for school size reported in Table 5.1 are likely to be especially sensitive to the weighting issue because schools were selected in the original HS&B design with probability proportional to size within a number of different strata. As a result, the corresponding school design weights strongly depend on school size (this is illustrated in Fig. 1). Consequently, if the weighting is not properly applied, the estimated effect of school size may be misleading.

On balance, Chubb and Moe would probably point us to their final analysis, presented in Table 5.7, of a school-level regression predicting effective school organization. In this analysis, school size does have a significant effect, but in the opposite direction (i.e. bigger schools appear better). Why and how this reversal of effect occurs, however, goes without comment.⁷ Such a reversal can only be caused by a relationship between school size and one or more of the other predictors included in the model. The key to sorting out these effects is to watch the school size coefficient as other predictors are individually entered. Presumably, the reversal occurs when the two bureaucratic influence

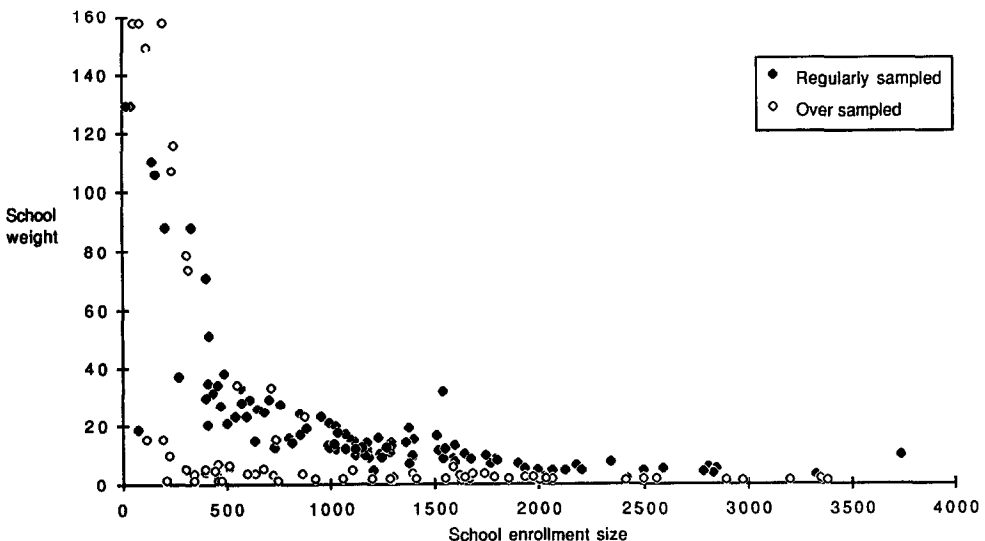


Figure 1. Distribution of school weights by school enrollment size, for regularly and oversampled schools in the Administrator Teacher Subset of High School and Beyond.

measures are added. If this speculation is correct, it would reinforce the argument that organizational size powerfully shapes work life within schools. It implies that the prevalence of bureaucratic constraints is related to the basic structure of the organizations to be managed — large public high schools — and not simply to the way they are governed. That is, the structure of individual schools and the norms that have built up around this structure drives, at least in part, the emergence of bureaucratic constraints.

Such results would lead us in a very different policy direction. We would focus much more attention on the big school as an organizational problem, and on districts where large schools are especially prevalent, e.g. our major cities. Smaller school units and greater decentralization of responsibility would appear as promising initiatives.

10. MARKETS OR ENHANCED DEMOCRATIC CONTROL: WHAT DOES THE EVIDENCE ACTUALLY SUPPORT?

Although it is obscured, Chubb and Moe actually provide supporting evidence that could be interpreted as supporting an argument for enhanced local control and decentralization, particularly in our large cities. They report in Table 5.11 that no parental attribute, including social class, has an independent effect on the prevalence of administrative and personnel constraints (their two measures of bureaucracy and lack of school autonomy). As a result, they state that “schools may be given the freedom to make their own decisions about personnel, curriculum, instruction and discipline even when their clienteles are not well educated” (p. 177). In further interpreting these results just four pages later, however, they contradict this evidence to conclude “market control tends to promote autonomy through its natural operation, while democratic control tends to allow for it only under the most favorable circumstances — outside of urban areas with **able and interested students and parents**”. Their evidence says that parental background does not affect school autonomy. Yet, when arguing for how to promote such autonomy, Chubb and Moe suddenly reverse directions and imply that it does.

Pursuing this a bit further, Chubb and Moe report in Table 5.12 that urban location has the biggest effect on the level of administrative and personnel

constraints, with private control the second most important (Table 5.11 also has these as the two most important predictors, although in the reverse order⁸). In fact, the urban location effect is probably underestimated because it is likely that this effect is almost exclusively in the public sector, resulting from differences between big, unresponsive urban school systems and client-sensitive suburban districts (there is no reason why the level of bureaucratic constraints in private schools should depend on urban location, because private schools governance arrangements are similar across locations).

We conjecture that if the urban location effect was estimated separately for public and private schools, the urban location effect for public schools would be much larger, probably dwarfing the private school effect. If this proved to be the case, it would appear then that bureaucratic constraints and ineffective school organization are principally an **urban public school** problem and not a more general democratic governance problem. Suddenly, the Chubb and Moe results begin to look more like an argument for radical decentralization of urban public schools in order to jolt bureaucratic calcification and revitalize democratic control!

Most troubling of all, in the final modeling of effective school organization, reported only a few pages later, where Chubb and Moe seek to establish the last component of their argument — that private control is the key influence on organizational effectiveness — the urban location variable is not even included as a potential predictor. Yet another technical decision was made that works strongly in favor of the authors’ argument. Chubb and Moe clearly have a personal preference for market control. Unfortunately, it is far from clear that the empirical evidence does.

11. SUMMARY AND CONCLUDING REMARKS

It is important that the reader understands the perspective we have brought to this critique. We have published considerable research on Catholic schools, most of which reflects quite positively on these schools. In our forthcoming book on this topic, we also speak quite favorably about school autonomy and how this can promote more effective institutions. Thus, in some important ways we cross-validate aspects of Chubb and Moe’s results and support at least components of their argument.

Nonetheless, we have taken a highly critical stance in this review. We have found *Politics, Markets, and America's Schools* lacking as a piece of disciplined inquiry on an important matter of public policy. Further, we judge its recommendations, although of some value, to be substantially overdrawn.

Disciplined Inquiry or Policy Argument?

Taken in total, the empirical argument traced out by Chubb and Moe has a hidden circularity. They begin by identifying high performing schools, but neglect to tell us that these schools are disproportionately private. To help their case, they doubly transform the student outcome metric, which further enhances the likelihood of private schools appearing in the high performing group. They then proceed to identify the organizational concomitants of high performance, which has now largely become a code word for private. Ignoring extant theory of school organization, they empirically pick the best components (best in the sense that they maximally predict high performance) to define effective organization, and then eventually lead us to the conclusion that private control is the key predictor of such organizations. Finally, when urban location emerges as a complicating factor, they conveniently ignore it. So they started with private, and low and behold ended up with private!

We also find troublesome the way this information was released to the public. The policy argument appeared in the *Public Interest* some 2 years before any scientific scrutiny was possible. Clearly this was going to be an important and controversial book which had the potential to substantially affect public policy. With the imprimatur of the Brookings Institution and its Washington, DC publication context, this potential was further heightened. It would seem reasonable to expect that careful scientific scrutiny would at least accompany the public policy debate. Yet, no statistical analyses were to be seen anywhere. Instead, the best various researchers could do was comment on the analytic plan described in the *Public Interest* article. It now appears that, based on these early criticisms, some further analyses were undertaken by Chubb and Moe to protect the core argument against criticism, but nowhere is this acknowledged. While the authors have striven for maximal public impact, it is far clearer that they have equally attended to the engagement of their scientific colleagues in debate over the merits and

interpretation of the evidence. This is a crafty policy argument, but it is not disciplined inquiry.⁹

The Evidence for Markets in Education

There is, in fact, considerable statistical evidence documenting internal organizational difference between public and private schools.¹⁰ The effective organizational practices, as discussed by Chubb and Moe, have been linked to greater school autonomy found in the private sector. However — and this is important to note — there is little empirical evidence that the effective organizational practices thought to accrue from greater autonomy actually foster student achievement.¹¹ Moreover, the validity of the evidence presented by Chubb and Moe on this point is doubtful for various reasons discussed above. While the lack of empirical evidence does not necessarily imply that schooling could not be improved through increased autonomy, it does suggest a closer look at the logic at work here.

Choice advocates argue that in a market system of control, competition will motivate schools to be more responsive to the needs of their current and potential clients. These arguments typically cite research documenting the effectiveness of private schools to support their claims. It is important to be very clear, however, about what this research shows. Many of the effective organizational practices identified by Chubb and Moe are actually more prevalent among non-Catholic than Catholic schools (Chubb and Moe, 1988). On the other hand, many of the positive outcomes described in the research comparing public and private schools are actually more typical of Catholic schools, which comprise about two-thirds of the ATS private school sample (Moles, 1988). For example, the more equitable social distribution of achievement, or “common school” effect (Coleman *et al.*, 1982), typifies Catholic schools but not other private schools. Similarly, the reduced dropout rates and unusual effectiveness of Catholic high schools for disadvantaged youth are not characteristic of private schools in general (Coleman and Hoffer, 1987). The specific effectiveness of Catholic girls’ schools (Lee and Bryk, 1986; Riordan, 1990) also does not appear to generalize across the private sector as a whole (Lee and Marks, 1991).

In sum, existing research clearly indicates differences among private schools in both their internal organizations and in the outcomes which accrue

from these organizational differences. This raises doubts about a blanket claim that a move toward greater privatization will accrue better consequences for students. As we have looked more closely at Catholic schools in an attempt to better understand how their organization might contribute to their unusual effectiveness, three key characteristics stand out. First is the constrained academic organization, discussed earlier, where common academic aims are held for all students regardless of background or future plans. The second characteristic focuses on the larger social organization in which the academic structure is embedded. Catholic schools tend toward a communal organization. They are smaller schools which promote more common experiences. The work of teachers is less specialized and more cooperative, and social relations are typified by informal and frequent face-to-face interactions among adults and students (the school life engendered in such communal organizations contrasts sharply with that found in the more bureaucratic structures typical of large, urban public secondary schools). Third, a distinctive set of shared

values animates schooling. Much in the daily life of Catholic schools symbolizes fundamental beliefs about individual dignity and a shared responsibility for advancing a just and caring society. An educational philosophy and practice aligned with these ideals helps promote social equity. We argue that this specific belief system, and not just any set of shared values, is required for the positive outcomes observed in these schools.

In sum, we must attend to the specific ideology at work in private schools if we are to understand the effects they produce. More attention must focus on the actual content of the values operative in schools, and the consequences linked to these values. In this regard, there is little evidence in our research on Catholic schools to substantiate the traditional micro-economic idea of human behavior motivated by entrepreneurial interest. While the competitive spirit, implied in a market system of incentives, may fuel economic development, it seems unlikely to engender the caring human relations and shared social responsibility so essential to institutions of human betterment.

NOTES

1. For a more extensive discussion of methodological weaknesses in Chubb and Moe see an earlier review by Witte (1990). For an insightful critique of the quality of their argument as policy analysis see Elmore (1991). We are indebted to both Witte and Elmore as they have helped to illuminate the path that we have followed in this review.
2. Recent school research has demonstrated that the exposure to subject matter is an important determinant of what students actually do learn. Further, these opportunities to learn various subjects are differentially afforded to students in different kinds of schools. As a result, school effects research has increasingly defined subject matter as a factor that must be included in the study design (McLaughlin *et al.*, 1990).
3. The actual analysis may be a bit more complex than just described. It appears that Chubb and Moe used the second transformation to create the gain scores but used the IRT measure for the pre-test. If this is correct, the relationship between the Chubb and Moe analysis and a simple ANCOVA becomes more complex. Nevertheless, our basic conclusion still holds.
4. Because of the complexities introduced here, serious questions were raised within the National Center for Educational Statistics as to whether adequate design weights could be even determined for this data subset. While arguably the ATS school weights provide a representative sample of schools, the HS&B student weights, employed by Chubb and Moe in their key student outcome analyses, appropriately apply only to the full HS&B data and not the ATS subset.
5. In fact, for the sample of public school students (Table D-4), the coefficient for the comprehensive measure of school organization (0.009) does not meet the conventional two-tailed standard of statistical significance, since the magnitude of the coefficient is not twice its standard error (0.005). Throughout their analyses, Chubb and Moe have used the more liberal one-tailed tests of statistical significance ($t > 1.64$ for $P < 0.05$). It is difficult to see how a combined set of directional hypotheses — the assumption of one-tailed tests — is actually justifiable in the analyses throughout the book.
6. For a further discussion of these results see a technical report by Bryk and Lee (1986) available by writing to the authors. Also, for a more general discussion in the context of multi-level models see Raudenbush and Bryk (1992).
7. A similar result, using these same data, was previously reported by Bryk and Driscoll (1988) in their study of the effects of school organization on teacher commitment and student engagement. Upon closer scrutiny of their analyses, Bryk and Driscoll surmise that two offsetting phenomena were at

- work: the greater prevalence of communal organization in small schools, and the enhanced opportunities for work specialization (for teachers to do their own thing) especially in relatively large suburban high schools.
8. We are puzzled by the differences reported in Tables 5.11 and 5.12 for the urban location and private control variables. Assuming both are dummy variables with the same coding scheme (e.g. a 1 indicating membership in the construct category and 0 elsewhere), then the percentile differences reported in Table 5.12 should be a simple monotonic transformation of the regression coefficients in Table 5.11. Curiously, however, they are not. The authors imply in the text that these discrepancies result from the unusual distribution of the outcome variable, but this distribution applies equally in transforming the effects of both the urban location and private control variables. It is not clear how this accounts for the differences reported in these two tables.
 9. For some reason, comparative research on public and private schools seems plagued by this problem. See an earlier essay review by Bryk (1981) for a further elaboration of this distinction between policy argument and disciplined inquiry.
 10. For an interpretative review of this evidence as it bears on high school organization and its effects, see Bryk *et al.* (1990).
 11. Bryk and Driscoll (1988) and Bryk and Lee (in press) demonstrate that the social organization of schools is strongly linked to teacher commitment and student engagement. Further, it is reasonable to believe that this greater willingness to sustain effort in more engaging institutions indirectly affects achievement [Bryk and Driscoll (1988) do provide some evidence on this point]. Ultimately, however, student achievement principally results from the interaction of students with teachers in classrooms around subject matter. Any adequate explanation of how organizational properties affect student achievement must postulate a mechanism that incorporates classroom life.

REFERENCES

- ASHTON, P.T. and WEBB, R.B. (1986) *Making a Difference: Teachers' Sense of Efficacy and Student Achievement*. New York: Longman.
- BIDWELL, C. (1965) The school as a formal organization. In *Handbook of Organizations* (Edited by MARCH, J.G.). Chicago: Rand McNally.
- BRYK, A.S. (1981) Disciplined inquiry or policy argument? *Harvard Educ. Rev.* **51**, 497–509.
- BRYK, A.S. and DRISCOLL, M.E. (1988) *The School as Community: Theoretical Foundations, Contextual Influences, and Consequences for Students and Teachers*. Madison, WI: National Center on Effective Secondary Schools, University of Wisconsin.
- BRYK, A.S. and LEE, V.E. (1986) *Design Weights and School Effects Research: A Cautioning Note*. Unpublished paper.
- BRYK, A.S. and LEE, V.E. (in press) *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- BRYK, A.S., LEE, V.E. and SMITH, J.B. (1990) High school organization and its effects on teachers and students: an interpretive summary of the research. In *Choice and Control in American Education*, Vol. 1 (Edited by CLUNE, W.H. and WITTE, J.F.), pp. 135–226. Philadelphia: Falmer.
- BRYK, A.S. and RAUDENBUSH, S.W. (1992) *Hierarchical Linear Models*. Newbury Park, CA: Sage Publications.
- BURSTEIN, L. (1980) The analysis of multi-level data in educational research and evaluation. *Res. Educ.* **8**, 158–223.
- CHUBB, J.E. and MOE, T.M. (1988) Politics, markets, and the organization of schools. *Am. Political Sci. Rev.* **82**, 1065–1087.
- COLEMAN, J.S. and HOFFER, T. (1987) *Public and Private High Schools: The Impact of Communities*. New York: Basic Books.
- COLEMAN, J.S., HOFFER, T. and KILGORE, S.B. (1982) *High School Achievement: Public, Catholic and Other Schools Compared*. New York: Basic Books.
- COOK, T.D. and CAMPBELL, D.T. (1979) *Quasi-Experimentation: Design and Analysis for Field Setting*. Boston: Houghton-Mifflin.
- DEWEY, J. (1966) *Democracy and Education*. New York: Macmillan.
- DEWEY, J. (1981) My pedagogic creed. In *The Philosophy of John Dewey* (Edited by McDermott, J.). Chicago: University of Chicago Press.
- DEWEY, J. (1900) *The School and Society*. Chicago: University of Chicago Press.
- DRISCOLL, M.E. (1989) *The School as Community*. Unpublished doctoral dissertation, University of Chicago.
- DURKHEIM, E. (1933) *Division of Labor in Society*. New York: The Free Press (1893).
- DURKHEIM, E. (1956) *Education and Sociology*. Glencoe, IL: The Free Press.

- DURKHEIM, E. (1961) *Moral Education*. Glencoe, IL: The Free Press.
- EASTON, J.Q., BRYK, A.S., DRISCOLL, M.E., KOTSAKIS, J.G., SEBRING, P.A. and VAN DER PLOEG, A.J. (1991) *Charing Reform: The Teachers' Turn*. Chicago, IL: Consortium on Chicago School Research.
- ELMORE, R.F. (1991) Book review of "Politics, Markets and America's Schools". *J. Policy Analysis and Management*, 687-695.
- FIRESTONE, W.A. and ROSENBLUM, S. (1989) Building a commitment in urban schools. *Educ. Evaluation and Policy Analysis* 10, 285-300.
- GAMORAN, A. (1987) The stratification of high school learning opportunities. *Sociol. Educ.* 60, 135-155.
- GOLDSTEIN, H.I. (1987) *Multilevel Models in Education and Social Research*. London: Oxford University Press.
- HAMBLETON, R.K. (1989) *Principles and selected applications of Item Response Theory*. In *Educational Measurement*, 3rd edn (Edited by LINN, R.L.). New York: Macmillan Publishing Company.
- HILL, P.T., FOSTER, G.E. and GENDLER, T. (1990) *High Schools with Character*. Santa Monica, CA: Rand Corporation.
- HOFFER, T., GREELEY, A.M. and COLEMAN, J.S. (1985) Achievement grown in public and Catholic schools. *Sociol. Educ.* 58, 74-97.
- LEE, V.E. and BRYK, A.S. (1986) Effects of single-sex secondary schools on student achievement and attitudes. *J. Educ. Psychol.* 78, 381-395.
- LEE, V.E. and BRYK, A.S. (1988) Curriculum tracking as mediating the social distribution of high school achievement. *Sociol. Educ.* 61, 190-94.
- LEE, V.E. and BRYK, A.S. (1989) A multilevel model of the social distribution of high school achievement. *Sociol. Educ.* 62, 172-192.
- LEE, V.E., DEDRICK, R.F. and SMITH, J.B. (1991) The effect of the social organization on teachers' efficacy and satisfaction. *Sociol. Educ.* 64, 190-208.
- LEE, V.E. and MARKS, H.M. (1991) Which works best? The comparative effectiveness of single-sex and coeducational independent secondary schools. Ann Arbor, MI: University of Michigan, School of Education. Unpublished paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.
- LORD, F.M. (1980) *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NH: Laurence Erlbaum Associates.
- MCLAUGHLIN, M.W., TALBERT, J.E. and BASCIA, N. (1990) *The Context of Teaching in Secondary Schools*. New York: Teachers' College Press.
- MASON, W.M., WONG, G.M. and ENTWISTLE, B. (1983) Contextual analysis through the multilevel linear model. In *Sociological Methodology*. San Francisco: Jossey Bass.
- MOLES, O. (1988) *High School and Beyond Administrator and Teacher Survey (1984)*. *Data File User's Manual*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- NEWMANN, F.M., RUTTER, R.A. and SMITH, J.S. (1989) Organizational factors affecting school sense of efficacy, community, and expectations. *Sociol. Educ.* 62, 221-238.
- PARSONS, T. (1959) The school class as a social system. *Harvard Educ. Rev.* 29, 297-318.
- POWELL, A.G., FARRAR, E. and COHEN, D.K. (1985) *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Boston: Houghton Mifflin.
- RAUDENBUSH, S.W. and BRYK, A.S. (1986) A hierarchical model for studying school effects. *Sociol. Educ.* 59, 1-17.
- RAUDENBUSH, S.W., ROWAN, B. and CHEONG, Y.F. (1992) Contextual effects on the self-perceived efficacy of high school teachers. *Sociol. Educ.* 62, 150-167.
- RAUDENBUSH, S.W. and WILLMS, J.D. (1991) *Pupils, Classrooms, and Schools: International Studies of Schooling from a Multilevel Perspective*. New York: Academic Press.
- RIORDAN, N.C. (1990) *Girls and Boys in Schools: Together or Separate?* New York: Teachers' College Press.
- ROSENHOLTZ, S.J. (1989) *Teachers' Workplace: The Social Organization of Schools*. New York: Longman.
- WALLER, W. (1932) *The Sociology of Teaching*. New York: Wiley.
- WEBER, M. (1947) *Theory of Social and Economic Organization*. New York: Macmillan.
- WITTE, J.F. (1990) Understanding high school achievement: After a decade of research, do we have any confident policy recommendation? Presented at the 1990 annual meeting of the American Political Science Association, San Francisco.