JOURNAL OF MANAGEMENT ACCOUNTING RESEARCH
Volume Twenty-One
2009
pp. 203–239

A Methodology for Evaluating the Cost-Effectiveness of Alternative Management Tools in Public-Sector Institutions: An Application to Public Education

Yaw M. Mensah
Michael P. Schoderbek
Robert H. Werner
Rutgers, The State University of New Jersey

ABSTRACT: The shift toward performance budgeting and outcome measures for public-sector institutions in recent decades has created a need to formally link inputs consumed and outcomes achieved. Given the inherent problems of cost accounting systems in public-sector institutions, we propose a statistical approach to identify the most cost-effective management tools that also recognize the endogeneity between costs and outcomes. The model developed allows for the examination of possible trade-offs that can be exercised by public-sector institutions facing multiple stakeholders with conflicting objectives. Using public schools in New Jersey and a set of variables identified from the education economics literature, we estimate cost and outcome functions to demonstrate empirically the choices made by school district superintendents that trade off the interests of various stakeholders, while seeking to meet the core objectives of the institutions. Our empirical results provide insight on the variables controllable by the superintendents that appear to be used inefficiently, or are subject to institutional constraints that limit the flexibility in input choice assumed by the proposed method. From a management accounting standpoint, the identification of such variables narrows the areas to be focused on in the search for improvements in performance.

INTRODUCTION

Over the past two decades, considerable attention has been focused on performance measurement and reporting for public-sector institutions. Much of this effort has been directed at moving from input-based control budgets to output- and outcome-based performance budgets. At the federal government level, the Government Performance and Results Act of 1993 initiated the most recent and comprehensive attempt at reforming the financial planning, management, and control system (Kravchuk and Schack 1996;

We gratefully acknowledge Yut'se Thomas and Allen DuPree of the New Jersey Department of Education for providing us with data used in this study. This study has benefited from funding by the Center for Governmental Accounting Education and Research (CGAER) of Rutgers University. This paper was presented at the Management Accounting Section of the American Accounting Association's 2007 Annual Meeting in Chicago. We thank the discussant (James Hesford) and the other conference participants for their helpful comments and suggestions. Finally, we are grateful to Joan Luft (editor) and the anonymous reviewer of this journal whose very insightful comments have contributed greatly in the revisions of this paper.

McNab and Melese 2003). Similar contemporaneous initiatives have been undertaken at the state and local government levels in states such as Texas, Oregon, Minnesota, Virginia, and Florida (Broom 1995).[1] These developments parallel the adoption by the Governmental Accounting Standards Board (GASB) of Concepts Statement No. 2, *Service Efforts and Accomplishments Reporting* in 1994 (GASB 1994), and subsequent GASB Standards No. 33 (GASB 1998) and No. 34 (GASB 1999). Yet, despite the comparative advantages of accounting researchers in knowledge of the measurement process (Kinney 2001), accounting research contributions to the public debate on how to increase the cost-effectiveness of public-sector institutions have been sparse, given the importance of that sector in national and state economies.

The move from input-oriented control budget systems to outcome-oriented performance budgets is motivated by several considerations, the most important of which are making public-sector institutions more efficient, more effective, and holding public managers accountable for policy outcomes. Control budgets can then be loosened so that public managers can more effectively employ their managerial discretion in the pursuit of desired outcomes (Gianakis 2002). For public managers to be more effective in the discharge of their duties, a systematic method of linking resource inputs to final outcomes on which management performance can be evaluated is needed. Although activity-based costing (ABC) has been touted in the literature as offering a means to link inputs used to the outputs of public-sector institutions (Mosso 1999; Brown et al. 1999; Mullins and Zorn 1999), the relationship of their inputs to eventual outcomes as perceived by the public cannot be defined in an *a priori* manner (McNab and Melese 2003). Establishing the *ex post* relationship between inputs and measured outcomes via statistical means, however, is feasible. Thus, assuming the temporal stability of the relationship between inputs and the final outcomes, public managers can develop a knowledge base of managerial tools that are cost-effective in achieving the desired outcomes.

At the same time, certain aspects of public-sector institutions are pertinent in understanding the constraints in which they operate, and consequently, the choices they can make among the feasible set of inputs and outputs. Hansmann (1996) has presented a theoretical framework for understanding organizational structures as one where the evolved structure minimizes the combined costs of ownership and of contracting. Hansmann (1996, 228–230) argues that non-profit institutions arise where the primary beneficiaries of the services provided are frequently in a poor position to determine, without significant effort or cost, the quality and/or quantity of services provided. As a result, assigning ownership of the institution to any other group of stakeholders would create severe agency problems (i.e., incentive and opportunity for the owners to exploit the beneficiaries). In addition, the cost to the beneficiaries to exercise effective control would be quite large relative to the value of their transactions with the institution. Thus, creating a non-profit institution allows the managers to operate the institution as fiduciary trustees for the intended beneficiaries. However, because there are multiple stakeholders, the managers must also be sensitive to the concerns of these other parties. Thus, the choices they can make in terms of the selection of inputs and managerial tools reflect multiple influences.

The purpose of this paper is to present empirical evidence of the trade-offs available to public-sector managers facing multiple objectives and conflicting demands. Because of the difficulty of measuring performance, public-sector institutions do not have high levels of bonus-based compensation. However, as noted by Rose-Ackerman (1996), the values,

[1] For a review of performance-based budgeting at the state level, see Willoughby and Melkers (1998, 2000).

motives, and incentives of managers of non-profit institutions may be less mercenary than those of for-profit organizations. Thus, even though incentive contracting can align the interests of the stakeholders and the public-sector managers, the incentive contracts must include non-pecuniary as well as pecuniary returns (Baber et al. 2002). By identifying some performance measures that can be chosen by public-sector managers under specific circumstances, and further differentiating between those contributing to cost efficiency and outcome effectiveness while simultaneously controlling for endogeneity among them, this study contributes to an understanding of how the effectiveness of control tools available to public-sector managers can be measured in multiple-objective situations.

The method we develop in this paper is applied to a sample of 521 New Jersey school districts. Two models are estimated simultaneously: a cost function (with per-pupil expenditures as the dependent variable) and an outcome function (with student achievement scores as the dependent variable). The two dependent variables are treated as jointly endogenous. Explanatory variables include those controllable by the school district superintendent (e.g., student-faculty ratio) and those that are uncontrollable (e.g., percent students receiving federal meal aid). The focus of analysis is on the controllable explanatory variables included in both the cost and outcome equations. By examining the statistical significance and signs (positive or negative) of the estimated coefficients, this paper develops a measure of *relative cost-effectiveness of the controllable variables.*

Consider the case where a particular variable is significant and positive in the outcome equation, but lacks significance in the cost equation. The positive coefficient suggests that a marginal increase in its use is associated with a marginal increase in test outcomes, but has no marginal (off-setting) effect on total costs. On the other hand, variables that are significant with positive signs in both the cost equation and the outcome equation indicate a trade-off where increased levels of that variable will increase both test scores and total costs. To evaluate these types of trade-offs, we compute a trade-off ratio from the estimated coefficients as a guide to school district administrators.

The paper is organized as follows. The next section presents the theoretical background and our methodology to evaluate the cost-effectiveness of the choices made by public-sector managers. The following section is a review of the educational economics literature on the determinants of school spending and achievement scores from which our choice of variables are based. Next, we provide the institutional setting of New Jersey's public school system and standardized tests. The following section includes specification of our cost and outcome functions. Afterwards, we present our empirical results, including an analysis of the relative cost-effectiveness of alternative management tools and contribution of our proposed methodology. A sensitivity analysis of our main results is provided in the next section, and the conclusions and recommendations follow.

THEORETICAL DEVELOPMENT AND PROPOSED METHODOLOGY

Theoretical Development

Theories of the economic foundation for non-profit institutions rest on the difficulty of measuring organizational performance in the settings in which they exist. For example, Hansmann (1996) argues that "public good" institutions arise in contexts where multidimensional criteria are needed to evaluate performance because of multiple stakeholders.

As an example of such "public good" institutions, consider the multiple stakeholders in the public education system: (1) the external stakeholders like the state legislature, state and local government officials, and taxpayers, and (2) the internal or beneficiary stakeholders like parents, school district administrators, and teachers. The state legislature has an interest in achieving high reported levels of educational outcomes subject to financing constraints. State and local government administrators are similarly concerned with outcomes, but the budgetary constraints are even more pressing since they must deal with local taxpayers and the specific financing of schools through such funding mechanisms as local property taxes. Taxpayers (whether through income taxes, sales taxes, or local property taxes) are very much focused on the tax burden imposed by the school system and thus on the efficiency with which the public schools are run.

In contrast to these external stakeholders, internal stakeholders such as parents are more concerned about the quality of the educational outcomes, and much less focused on the budgetary aspects. Teachers, on the other hand, are expected to be equally concerned about their personal remuneration, the quality of students, the instructional resources provided, and the ultimate educational outcomes as measured by objective measures such as statewide test scores. School district administrators therefore form the apex of these convergent forces where a balance must be maintained between the external and internal stakeholders. It is worthwhile to note that the very key role played by district superintendents in moving urban school districts forward in terms of measurable achievements required under the federal No Child Left Behind Act of 2001 (U.S. Department of Education 2002) (under the auspices of Harvard University's Public Education Leadership Project) has been observed by Childress et al. (2006).

Three other pertinent aspects of the non-profit (including the public sector) environment are: (1) the lack of well-defined production functions that tie inputs with outcomes; (2) the frequent inability to measure outcomes quantitatively or over a short-term horizon; and (3) the tendency of some stakeholders to focus on the more easily measured quantitative outcomes with more limited attention to outcome quality.[2] Our objective in this paper is to contribute to the first area by presenting a methodology that ties inputs with outcomes (or reasonable proxies for difficult-to-measure outcomes.) Our objective is thus similar to Dopuch and Gupta (1997), except that we seek to exploit the specific peculiarity of public-sector and non-profit institutions—the lack of a well-defined production or associated cost function. Thus, in contrast to the single equation approach adopted by Dopuch and Gupta (1997), we use a two-equation approach. In addition, instead of their stochastic frontier estimation approach, we examine the use of both seemingly unrelated regressions and simultaneous equations approaches. Our approach is thus focused on identifying the average behavior of public-sector administrators engaged in actions designed to satisfy multiple constituencies, rather than an approach that assumes that such administrators are necessarily concerned with cost-minimization or output-maximization.[3]

Proposed Methodology

For public-sector institutions, a key resource constraint that managers face is the level of funds available for expenditure purposes. Given that the funds available are typically appropriated by legislatures and the services provided are frequently either free or at less

[2] As examples of this third aspect, the *output* of hospitals is often measured in terms of the number of surgical procedures performed, number of patients treated, total patient days, and number of outpatient visits. However, their final *outcome* measures focus on mortality rates, patient recovery rates, and longevity of patients after treatment. Similarly, for correctional institutions, *output* measures tend to focus on the total number of inmates and days of incarceration, while final *outcome* measures focus on the effectiveness of incarceration as captured by the rate of recidivism (Mensah and Li 1993).

[3] Specifically, we estimate a cost function using the inputs as traditionally defined, and also estimate an outcome function using most of the same explanatory variables as in the cost function. Thus, from the estimated coefficients in the cost and outcome equations, we are able to derive directly comparable measures of the marginal cost and marginal benefits of the input factors.

than full cost to the users, the effective management of expenditures is of paramount importance. At the same time, the funding agency or legislature, as well as taxpayers, are well-attuned to perceptions of the quality of services delivered by the institution. These countervailing forces suggest that managers of public-sector institutions must necessarily be concerned with their expenditure levels and the outcomes they deliver to their stakeholders.

Managers of public-sector institutions face a multi-dimensional objective function, but as noted above, the two key variables are the cost of providing the services and the outcomes they are able to achieve. We propose a methodology that ties together a possible set of management control tools to assist in management's decision that also deals with the potential endogeneity inherent in this setting. Consider a simple two-equation system given by:

$$Y_1 = a_0 + (a_1X_1 + ... + a_nX_n + a_{n+1}Y_2) + (\lambda_1Z_1 + ... \lambda_mZ_m), \tag{1}$$

and:

$$Y_2 = c_0 + (b_1W_1 + ... + b_sW_s + b_{s+1}Y_1) + (\gamma_1Z_1 + ... \gamma_mZ_m), \tag{2}$$

where:

Y_1 = cost;
Y_2 = outcome;
$X_1 \ldots X_n$ = exogenous variables that are theorized to drive cost levels;
$W_1 \ldots W_s$ = exogenous variables that are theorized to determine outcome levels; and
$Z_1 \ldots Z_m$ = management decision variables that may affect both costs and outcome levels.

In Equations (1) and (2), the variables denoted X and W could be the same variables, but they differ from the Z variables in that they are presumed to be truly exogenous in that they are not subject to management or administrative control. So X and W could be input prices (where the institution is a price-taker for that input), socio-economic variables that drive either costs or demand for the public services provided, or other exogenous variables. As long as the set of X variables are not the same as the W variables, the simultaneity implied by the two system of equations can be resolved since the variables unique in either equation would serve as the instruments for the dependent variable in that equation.

The two-system of equations can be estimated using ordinary least-squares (OLS) on the reduced form equations, jointly using seemingly unrelated regressions (SUR), or simultaneously using either two-stage least-squares (2SLS) or three-stage least-squares (3SLS). The endogeneity presumption embodied in the system of equations can be evaluated in a three-step process:

(1) a test of the fit of the instruments to deal with the weak instruments problem (Staiger and Stock 1997; Stock et al. 2002);
(2) a test of over-identified restrictions if either of the two equations is overidentified (Bound et al. 1995; Hahn and Hausman 2003); and
(3) a Hausman (1978) test of the null hypothesis that no endogeneity problem exists (1978).

The three-step process described above is a standard econometric technique for estimating simultaneous equations. The innovation in this paper is in relating the coefficients in the cost equation to the coefficients in the outcome equation in a systematic manner to derive the cost-effectiveness of the various controllable variables. In the standard neoclassical production setting, the normal approach is to estimate either a cost function (where cost is the dependent variable) or a production function (where output is the dependent variable). Estimating both functions together in a system of equations is never done in practice to our knowledge because, in such setting, the cost function is the dual of the production function, so no new information can be gleaned from such a system of equations in which both the cost and production functions are estimated together (see, for example, Shephard 1970; Varian 1992, 83; Mundlak 1996). What is unique about the public-sector setting is that the lack of a well-defined production technology defies the closed production system implied in neoclassical production settings where such a unique transformation function between the cost and production functions exist.[4] Thus, although cost and production functions have been estimated previously in the educational economics literature to describe the public school education system, this paper is the first to estimate both jointly and to explore the resulting coefficients to provide insight into the cost-effectiveness of the controllable variables.

Assuming the econometric problems in estimating Equations (1) and (2) are resolved, the relationship between λ_I and γ_ι provides the primary focus for analysis. Note that Equation (1) is a cost function, so a positive sign for λ_I implies that a factor is cost-increasing. At the same time, Equation (2) is an outcome function, so that a positive sign for γ_ι indicates that variable is effective in increasing outcome levels.

From a management standpoint, the most interesting insights are gained by comparing the statistical significance of variables across the cost and outcome equations. Variables that are statistically significant with positive signs in both the cost and test outcome equations signify that the tools, though effective, are also costly. Thus, a trade-off is involved between cost and outcomes. To help in identifying such variables, the following labels will be used in the empirical section of the paper:

BC = variable that is beneficial to costs, but has no effect on outcomes;
BO = variable that is beneficial to outcomes, but has no effect on costs;
DB = variable that is doubly beneficial (i.e., beneficial effects on both costs and outcomes);
DN = variable that is doubly negative (i.e., negative effect on both costs and outcomes);
CI = variable that has an increasing effect on costs, but no effect on outcomes;
NO = variable that has a negative effect on outcomes, but no effect on costs; and
TO = variable with a trade-off effect (i.e., either a beneficial [negative] effect on costs offset by a negative effect on outcomes, or vice versa).

In general, variables labeled DB are the most effective, followed by BC and BO variables. TO variables are useful policy tools, but the costs and effects must be weighed

[4] In this paper we define a "closed production system" as one in which all inputs and outputs are known and measurable and included in the equation estimated. For example, in the standard neoclassical system, the typical inputs of labor, materials, and capital are known and measurable, just as the outputs of the system are. This results in a closed system. The inclusion of price information allows the estimation of a cost function that is the dual of the production function. In the public-sector setting, the production system is an open system. Beyond the labor, supplies, and capital inputs, the environmental, socio-economic, and political settings can greatly influence the outcomes, which they themselves are not easily quantified.

carefully. Finally, DN variables in particular (but also CI and NO variables) are to be avoided if possible, or minimized in their use.

For variables given a rating of TO, the ratio $\pm|\gamma_i|/|\lambda_i|$ (i.e., the absolute values of the outcome equation coefficient to the cost equation coefficient) is derived and labeled *an index of relative cost-effectiveness*. This index is given a sign ($\pm$) that corresponds to the sign that appeared in the coefficients for the cost and outcome equations (since TO variables have the same sign in both equations). Higher values for the index with positive signs signify that higher outcome is achieved for each dollar of spending (with due allowance made for the measurement units of the variable under consideration). In contrast, indices with negative signs signify that outcomes are sacrificed for each cost dollar injected. That is, when the index value has a negative sign, the variable in question lowers both costs and outcomes. Thus, the higher the absolute value of the cost relative to the outcome, the more cost-effective that variable is (in the sense of a trade-off). The reverse is true for indices with positive signs.

The models presented in Equations (1) and (2) also permit the decisions made by public-sector managers to be evaluated in an *ex post* setting. Specifically, assuming the incentive contracts under which they are operating take into account the differing (and potentially conflicting) objectives of the multiple stakeholders, the model proposed permits the resulting pattern of choices to be analyzed. Such an analysis would permit the costliness of the constraints faced by public-sector managers to be better understood. For example, the consistent appearance of CI (cost increasing) and NO (no benefit to outcomes) controllable management tools would imply the existence of institutional factors that limit the managers' ability to influence the usage of those tools. Such constraints may exist if legislatures, donors, or the public, in well-meaning but misguided efforts, impose constraints to the ability of the managers to change the mix of inputs.[5]

Research Questions

As an empirical demonstration of the proposed method, we apply our methodology to the public school system in the State of New Jersey. Public schools dominate private schools in relative numbers of students, and education represents a very important public-sector institution. Continuing concerns exist on the cost-effectiveness of public education, with frequent calls for reform of the education system. Thus, the following research questions are investigated using our methodology:

(1) What factors are posited to influence the levels of school expenditures and test score outcomes in the public schools districts? Which of these factors can be deemed controllable by the school district superintendent?
(2) Which of these factors are common and which are unique in their influence on the two dependent variables (per-pupil expenditures and student achievement scores)?
(3) Given the signs and magnitudes of the coefficients on the controllable variables in the cost and test outcome equations, which factors are cost-effective, cost-ineffective, or involve a trade-off between costs and benefits.

To investigate these questions, the next section reviews the public education literature to identify the X_i, W_i, and Z_i variables specified in our models. Subsequently, the institutional

[5] See, for example, Mensah and Li (1993) on the effects of constraints such as line-item budgeting on allocative efficiency.

setting of the New Jersey public schools are described, and Equations (1) and (2) are operationalized.

PRIOR LITERATURE ON PERFORMANCE MEASUREMENT IN PUBLIC SCHOOLS

Since the Coleman report (Coleman et al. 1966),[6] the literature on performance measurement in public schools has largely focused on the effect of school inputs on student achievement scores. This research, centered on the ''Does Money Matter'' debate, is relevant for the selection of variables for our outcome function. However, it should be noted that variables posited to influence test scores (e.g., the teacher-student ratio), more often than not, affect school expenditures as well. Thus, most of the variables discussed are relevant for our cost function as well as our test outcome function. It is this endogeneity between school inputs and test outcomes that motivates our attempt to develop a methodology to analyze the beneficial trade-offs between the two. The predicted influence on costs and outcomes of the variables identified below is discussed later in model specification.

Hanushek (1986) surveyed research from 38 different articles that utilized an educational production function approach to explain standardized test scores. A set of core explanatory variables were included in these production functions, including the teacher-pupil ratio, teacher education, teacher experience, teacher salary, and expenditures per pupil (Hanushek 1986, 1161). Although Hanushek (1986, 1162) ultimately concluded: ''There appears to be no strong or systematic relationship between school expenditures and student performance,''[7] combinations of these variables have been widely used in the production functions of nearly all studies on the school expenditures-test score debate. Thus, these core explanatory variables, with the exception of teacher education,[8] are used in our models as controllable factors to explain achievement scores.

In addition to these teacher-related factors, more recent studies have also included administrative-related factors in their production functions. Brewer (1996), who uses number of school district personnel by functional area to explain test scores, includes separate regressors for district administrators and building administrators. Jaggia and Kelly-Hawke (1999) include per-pupil administration expenditures and total per-pupil expenditures as determinants of test scores, whereas Dee's (2005) model includes instructional expenditures per pupil and non-instructional expenditures per pupil as explanatory variables. To capture resources devoted to administration, we include the median administrative salaries and the median years of administrator's experience. These two variables, both controllable by the school district, parallel the construction of our two teacher-related variables: median faculty salaries and median years of faculty experience. Additionally, to examine the effect of resources deployed on instruction *vis-à-vis* administration, we include the relative cost shares for instruction (instructional expenditures/total expenditures) and administration (administrative expenditures/total expenditures) in our models. It should be noted that, by

[6] The Coleman report, based on a database of 3,000 schools and over a half million students, concluded that cross-sectional differences in family background and student characteristics were the primary determinants of student achievement, not school inputs (e.g., per-pupil expenditures).

[7] Hedges et al. (1994; hereafter HL&G), who criticize Hanushek's (1986, 1989) ''vote-counting'' methodology, reanalyzed Hanushek's data using a meta-analysis approach. The results of HL&G's (1994) combined significance tests find a systematic positive relation between school resource inputs and student achievement tests.

[8] Of the 106 educational production functions summarized by Hanushek (1986, 1161), teacher education was positively related to test scores in only 6 percent of the models, and negatively related in 5 percent of the models.

using relative input cost shares, we mitigate potential multicollinearity problems from deflating, for example, both instructional expenditures and administrative expenditures by number of students.[9]

We include two additional controllable factors in our model, the number of students per computer and the student attendance rate. The number of students per computer proxies for the degree to which modern instructional technologies have been introduced into the classroom. Elliott (1998) used teacher survey data to construct a variable for computer usage. Student attendance rate, while not used as a regressor in the studies we surveyed, is identified as an indicator of efficiency by the GASB (1989) in its research report, *Service Efforts and Accomplishments: Elementary and Secondary Education*.

Myriad variables have been used in the literature to control for the effect of socio-economic factors on achievement scores. Sander (1993) includes family income and three indicator variables (Black, Hispanic, and Poor) in his model to explain scores on the ACT college entrance exams. In their model to explain Massachusetts test scores, Jaggia and Kelly-Hawke (1999) include four variables to capture students' family background: percent rental units within the community, percent single mothers, community crime rate, and percent professionals and managers living in the community. Dee (2005), who explains high-school graduation rates, uses five variables to control for socio-economic priors: percent children at risk, percent children who speak English "not well" or "not at all," percent householders with high-school degrees, percent householders with some college, and median income in households with children. Elliott (1998), who uses hierarchical linear modeling to explain math and science achievement scores, includes indicator variables for gender, minority status, and urban school; and continuous variables for socio-economic status (constructed from parent's education, occupation, and income), percent students receiving free lunch, percent students in special education programs, and percent students with limited English proficiency.

It should be noted that many of the variables used above capture similar dimensions of socio-economic background. The studies surveyed and variables utilized in these studies are not intended to be all-inclusive. The selection of socio-economic factors used in empirical studies on the school input-output relation is often driven by data availability, and this current study is no exception. The variables we used to capture socio-economic factors are discussed in the section in model specification.

In contrast to the number of studies that examine the effect of school inputs on student outcomes, relatively few papers have modeled educational cost functions. Dopuch and Gupta (1997; hereafter D&G), used stochastic frontier estimation (SFE) to examine the efficiency in providing education for 446 Missouri school districts. D&G (1997) model total school expenditures as a function of number of students, standardized test scores, percentage of aid dependent or orphan students, percentage of minority students, percentage of students continuing on to higher education, and average income level. OLS regression estimation found all independent variables to be significant in the hypothesized direction in explaining school expenditures.

Like D&G (1997), we specify a cost function to investigate the determinants of per-pupil spending. But our methodology also employs a production (outcome) function to explain test scores, like the earlier studies synthesized by Hanushek (1986). We then use

[9] Jaggia and Kelly-Hawke (1999, 193) acknowledge that mutilcollinearity may exist among their school input variables since per-pupil expenditures are mostly a function of the teacher-pupil ratio and per-pupil administrative expenditures.

the estimated coefficients from these models to analyze the most cost-effective management tools by school district superintendents in providing education in the presence of multiple stakeholders.

INSTITUTIONAL SETTING

New Jersey School Districts

In New Jersey, the school system is highly decentralized with multiple levels of controls. At the top of the hierarchy, there is a county-wide superintendent who oversees all the schools in the county (of which there are 21 in the state). Within the county are the local school boards and the school district superintendents. Each locality or municipality has its own school districts, although some regional school districts (particularly at the high-school level) also exist. The focus of interest in this study is the local school district superintendents.

The state has five distinct types of school districts: elementary school districts (Kindergarten to Grade 6 = *ELEM*), middle school districts (Kindergarten to Grade 8 = *MIDD*), K–12 school districts (Kindergarten to Grade 12 = *K12*), high-school districts (Grades 7 to 12 and Grades 9 to 12 = *HIGH*), and county vocational high-school districts (Grades 8 to 12). Of the total 572 school districts in the state, there are 66 *ELEM*, 223 *MIDD*, 215 *K12*, 47 *HIGH*, and 21 vocational school districts. Because the vocational schools are quite different from the other four districts in terms of their emphasis on occupational as opposed to academic achievement, this study excludes the vocational school districts.

The school district superintendent who oversees the districts operations is hired by the Board Of Education on a multi-year contract, typically three to five years. The superintendent hires additional administrative personnel at the district-level, such as Business Administrators, Directors of Curriculum, Directors of Human Resources, etc., depending on the type of district (e.g., K–12, K–8) and size of the district. School principals are also hired on multi-year contracts by the district superintendent, with consultation by the Board of Education, and are granted tenure after three years of service. Teachers are hired on renewable annual contracts by the principals, with consultation by the District Superintendent, and also earn tenure after three years. Teachers in New Jersey are members of the American Federation of Teachers (AFT), and operate under union rules. Thus, seniority is a large component of teachers' salary increases, although principals and school district superintendents may also reward teachers for superior performance in the classroom.

In controlling instructional costs, district superintendents have a wide array of tools at their disposal, notwithstanding teacher tenure and seniority pay scales. According to the New Jersey Department of Education (2004; hereafter NJDOE), approximately one-forth of New Jersey's full-time teachers have zero to three years of teaching experience and are thus not tenured, nor are part-time teachers that comprise 2.2 percent of the teaching workforce. There is also wide flexibility in the use of teacher aids (non-tenure-track positions), which number one-fourth the number of full-time classroom teachers (NJDOE 2004), and educational support service personnel, that include teachers with special skills for which state certification is required (handicap learning specialists). In controlling costs Superintendents can shift teaching personnel to various assignments, depending on their educational needs and cost objectives. As an example of a cost-reduction mechanism, enrichment programs or early intervention programs can be curtailed or eliminated, and the teaching staff from these programs can be re-assigned to full classroom duties to replace retired teachers. Thus, in this study instructional costs as well as administrative costs are treated as controllable by the school districts.

District Funding Mechanisms

Public schools in New Jersey are funded almost exclusively through property taxes levied at the school district level. The exception to this rule is the school districts designated as Abbott districts. Abbott school districts have been identified by the New Jersey State Supreme Court (or through subsequent legislative processes) as "poorer urban districts" with both "poverty and educational inadequacy" so substantial that "poorer disadvantaged students" cannot compete with "relatively advantaged students" (Librera 2003). For these 30 Abbott districts (all K–12), the state provides "Abbott parity aid" that is calculated to provide them with the same per-pupil operating budget as would be found in New Jersey's wealthiest school districts. In addition, under the New Jersey State Supreme Court Abbott decisions, school districts that show financial inability but do not meet the Abbott criterion are eligible for Discretionary Education Opportunity Aid.[10]

Standardized Achievement Tests and Performance Measurement

Public school students are required to sit for three sets of standardized tests given in Grade 4, Grade 7, and Grade 11 (shifted to Grade 12 in 2002). Since our interest is on capturing a single measure of outcomes across all four types of school districts, we adopt an outcome measure based on Callan and Santerre (1990) and D&G (1997). School district output is defined as the product of quantity of output (total student enrollments = *ENROLL*) and quality of output (measured by test score outcomes = *CTEST*).[11] The composite test score (*CTEST*) for the jth school district is constructed as follows:

$$CTEST_j = \frac{1}{\sum_{i}^{K_C} S_{ij}} \sum_{i}^{K_C} S_{ij}^{*} \left[TEST_{ij}/TEST_{i\mathrm{MAX}}\right], \tag{3}$$

where:

S_{ij} = the number of students in the jth district taking $TEST_i$;
$TEST_i$ = the standardized state-administered test given in Grades 4, 7, and 11–12;
$TEST_{i\mathrm{MAX}}$ = the maximum score for $TEST_i$, attained by any district for that year; and
K_c = 1 for *ELEM* and 9–12 *HIGH* school districts, 2 for *MIDD* and 8–12 *HIGH* school districts; and 3 for K–12 school districts.

As calculated above, *CTEST* is the weighted average percentage score (relative to the maximum score attained in that year), where the weights are the total number of students who took that test in that district. Scaling by the maximum score attained in that year was necessary because the maximum scores differ across grade levels, and lack of scaling would have biased the composite score in favor of districts that did comparatively better on the tests with higher maximum scores.

The use of standardized test scores as the sole measure of school district performance has been criticized as inadequate since public schools pursue multiple objectives (Hanushek

[10] A legislative history of the *Abbott vs. Burke* court cases and the criteria employed by the New Jersey Commissioner of Education in the designation of Abbott Districts can be found at http://www.nj.gov/njded/abbotts/regs/criteria.htm.

[11] The Grade 11 exam was replaced in 2002 by a different exam given in Grade 12. For all three tests, the basis skills evaluated during the period 2000 to 2002 are language and mathematics.

1979, 1986). Alternative measures employed in the literature are the dropout rate, high-school graduation rates, the college placement rate, and post-graduation earnings. However, measures based on eventual outcomes are applicable only to the K–12 and *HIGH* districts, and not applicable to this study since we pooled data from all four types of school districts. Moreover, the state of New Jersey itself uses the results of the standardized tests as a summary measure of the public educational process, and the public accountability reports issued for the schools emphasize these results almost exclusively. Thus, as a basis for contracting, this composite test score measure is quite appropriate.

EDUCATIONAL COST AND OUTCOME FUNCTIONS

Specification of the Cost Function

To address the research questions above, we operationalize the model in Equation (1) by identifying the factors that are controllable and uncontrollable by the school district superintendents. The seven uncontrollable factors in our cost function include some of the main variables found to affect costs or outcomes in prior studies. The uncontrollable factors are: (1) weighted district factor grouping index (*WDFG*); (2) geographic cost index (*GEOCEI*); (3) an indicator variable for Abbott districts (*ABBOT*); (4) the proportion of students in the school district *not* receiving federal meal aid (*HIGHINC*); (5) school size (measured as the natural log of the student population = *ENROLL*); (6) the number of students enrolled in special education programs (*SPED*); and (7) the number of students in limited English proficiency programs (*LEP*).

The first two variables are from year 2000 Census data and are constant for all three years. The New Jersey Department of Education used factor analysis to construct a District Factor Group index that reflects socio-economic factors linked in the educational literature to test score performance.[12] Based on the factor scores, the state classified the school districts into one of eight ordinal rankings, with one denoting the most disadvantaged, and eight the most privileged. We use the same ordinal rankings for the variable *WDFG*.[13] The second variable, *GEOCEI*, captures geographic differences in the cost of living across the state that impact on the cost of providing education. Both *WDFG* and *GEOCEI* are expected to be positively related to costs in our cost equation.

Because the Abbott mandate created by the New Jersey Supreme Court may create a lesser incentive for cost efficiency, the coefficient on the dummy variable, *ABBOT*, is expected to be positive. The variable *HIGHINC* is computed as the complement of the percentage of the student body from poor families who qualify for federal food aid.[14] High values (capped at 100 percent) reflect a concentration of higher-income families in the district. In general, such school districts would be expected to spend higher amounts per pupil because they can raise more property taxes to finance the school district.[15] However, because the state of New Jersey attempts to equalize spending by directing more funds to needy districts, the observed spending per pupil may be more reflective of cost efficiency.

[12] The factors used in the *NJDOE* factor analysis are (1) the number of single mothers in the school district; (2) the average income level in the district; and (3) the concentration of disadvantaged minorities in the school district.

[13] The exact variables included in the factor analysis program and the factor loadings are considered proprietary by the state, so we could not obtain more detailed information.

[14] The complement of *LOWINC* (percentage of students whose families qualify for federal food aid) is used in order to avoid dealing with negative values in the subsequent log transformation of the variable.

[15] A well-known economic theory offered by Tiebout (1956) posits that spending on government services will vary widely as a function of differences in wealth. Thus, wealthier areas can spend more from local resources on schools, and although court mandates and public policy may seek to equalize such spending across districts, differences may still survive.

Total student enrollments (*ENROLL*) are expected to have the most significant effect on total expenditures. As shown later, when the cost function is defined in terms of spending per pupil, *ENROLL* is expected to have a negative sign, reflecting the existence of economies of scale. In addition, to allow for possible diseconomies of scale, the square of the natural log of *ENROLL* is included in the model. This squared term is expected to have a positive sign if diseconomies of scale exist (D&G 1997; Greene 1980). Students enrolled in special education programs or in language proficiency programs can substantially increase the cost of educating students, so positive signs are predicted on the variables *SPED* and *LEP*.

The controllable factors (i.e., controllable by the district superintendent) included in the model are more often employed in production functions to explain test scores (see ''Prior Literature on Performance Measurement in Public School'') but are also posited in this study to influence costs. These controllable factors (grouped into student-related, faculty-related, and administrative factors) are summarized below:

(1) Student-Related Factors: These factors consist of the student-faculty ratio (*STUFAC*), the number of students per computer (*STUCOMP*), and the student attendance rate (*ATTD*).
(2) Faculty-Related Factors: These factors consist of the median level of faculty salaries (*FACSAL*), the ratio of instructional expenses to total expenditures (*CSINS*), and the median years of faculty experience (*FACEXPR*).
(3) Administrative-Related Factors: These factors consist of the median level of administrative salaries (*ADMSAL*), the ratio of administrative costs to total operating expenditures (*CSADM*), and the median level of administrators experience in years (*ADMEXPR*).

STUFAC reflects the average number of students per teacher in the school district. Since higher ratios indicate larger class sizes, the coefficient of *STUFAC* is expected to be negative in the cost function. Higher ratios of *STUCOMP* reflect lower degrees of penetration of modern instructional technology. We expect the coefficient for *STUCOMP* to be negative in the cost equation, reflecting the expectation that the introduction of modern instructional technology is costly. The coefficient on *ATTD* (student class-attendance rate) in the cost equation is expected to be negative, because as school district's fixed costs are spread over higher volumes, costs per pupil should decrease.

FACSAL is predicted to be positive in the cost function because higher median salaries for the teachers should translate into higher overall per-pupil spending. No prediction is made on the sign of the cost share for instructional spending (*CSINS*), because it is difficult to determine *ex ante* whether greater resources spent on instruction (relative to the other functional cost areas) will lead to higher or lower overall spending.[16] The median years of teaching experience (*FACEXP*) is expected to have a positive coefficient in the cost equation, reflecting the expectation that more seasoned teachers will earn higher pay than teachers with less experience.

For administrative-related factors, *ADMSAL* is expected to exert an upward pressure on spending, leading to an expected positive coefficient. There are two possible ways to view the cost share for administrative spending (*CSADM*). One view is that the more highly paid

[16] The five major functional cost classifications adopted by the NJDOE are Instructional, Administration, Student Support, Operations and Maintenance, and Extra-Curricular Activities, and are derived from *Financial Accounting for Local and State School Systems* (U.S. Department of Education, Office of Educational Research and Improvement 1990).

and hence more competent the administrators are, the greater the degree to which cost efficiency can be achieved, leading to a negative coefficient in the cost equation. The alternative view is that *CSADM* should have a positive sign, reflecting the frequently held view that the ratio of administrative costs to total spending is a measure of inefficiency in not-for-profit settings. The expected sign of the final administration-related variable, *ADMEXPR*, is also indeterminate in the cost function, because there is no strong *a priori* reason to expect that *ADMEXPR* will necessarily lead to lower overall costs.

In addition to these uncontrollable and controllable factors, the quality outcome measure *CTEST* may be considered as an explanatory (potential jointly endogenous) variable in the cost equation. Communities not satisfied with the level of student achievement in their public schools are likely to be motivated to adopt measures (including increased spending via pressure on the state government and increased local taxes) to raise test scores. Thus, a positive coefficient is predicted on *CTEST*.

Based on the foregoing, and using D&G (1997) as a guide for the functional form, the complete cost equation can be presented as:

$$\begin{aligned} TCOST_j = {} & \alpha_0 ELEM^{\alpha_1} MDD^{\alpha_2} HIGH^{\alpha_3} CTETS^{\alpha_4} ENROLL^{\alpha_5} (ENROLL^{\ln ENROLL})^{\alpha_6} \\ & WDFG^{\alpha_7} GEOCEI^{\alpha_8} SPED^{\alpha_9} LEP^{\alpha_{10}} HIGHINC^{\alpha_{11}} ABBOT^{\alpha_{12}} \\ & STUFAC^{\lambda_1} STUCOMP^{\lambda_2} ATTD^{\lambda_3} FACSAL^{\lambda_4} CSINS^{\lambda_5} FACEXPR^{\lambda_6} \\ & ADMSAL^{\lambda_7} CSADM^{\lambda_8} ADMEXPR^{\lambda_9} e_j \end{aligned} \tag{4}$$

After dividing through both sides of Equation (4) by *ENROLL*, and taking the natural log of both sides, the average cost function estimated can be written as:

$$\begin{aligned} \ln(EXPPP) = {} & \alpha_0 + \alpha_1 ELEM + \alpha_2 MIDD + \alpha_3 HIGH + \alpha_4 \ln(CTEST) \\ & + (1 - \alpha_5)\ln(ENROLL) + \alpha_6(SQ\_ENROLL) + \alpha_7 \ln(WDFG) \\ & + \alpha_8 \ln(GEOCEI) + \alpha_9 \ln(SPED) + \alpha_{10} \ln(LEP) \\ & + \alpha_{11} \ln(HIGHINC) + \alpha_{12} ABBOT + \lambda_1 \ln(STUFAC) \\ & + \lambda_2 \ln(STUCOMP) + \lambda_3 \ln(ATTD) + \lambda_4 \ln(\mathrm{FACSAL}) \\ & + \lambda_5 \ln(CSINS) + \lambda_6 \ln(FACEXPR) + \lambda_7 \ln(ADMSAL) \\ & + \lambda_8 \ln(CSADM) + \lambda_9 \ln(ADMEXPR) + e \end{aligned} \tag{5}$$

where:

EXPPP = total expenditures per pupil; and
SQ_ENROLL = [In (*ENROLL*)].^{2}

Note that the coefficient for *ENROLL* $(1 - \alpha_5)$ would be negative if there are economies of scale in the initial ranges of school district size. A significant positive value for *SQ_ENROLL* is expected if diseconomies of scale exist within the range.[17]

Specification of the Outcome Function

As noted earlier, the output of public schools is a product of quantity of students and outcome quality. In the operationalization of our outcome function in Equation (2), we

[17] See D&G (1997) for the optimal size formula.

include the potential jointly endogenous variable *EXPPP* (expenditures per pupil) as a determinant of test scores. Given the previous (though inconsistent) findings in the literature (Hanushek 1986; Hedges et al. 1994), we expect *EXPPP* to have a positive coefficient. Estimating a quality outcome function that includes *EXPPP* as an explanatory variable introduces mutual simultaneity into the relationship. The selection of explanatory variables must consider the simultaneous equation identification issue, and the need for suitable instrumental variables.

Six of the socioeconomic variables included in the cost equation are also included in the outcome equation: *SPED*, *LEP*, *HIGHINC*, *ABBOT*, *WDFG*, and *GEOCEI*. *SPED* is expected to have a negative coefficient in the outcome function because it is a handicap in the learning process. The expected sign of *LEP* is uncertain. Although limited English proficiency may be an initial handicap, effective teaching can overcome this initial disadvantage. *HIGHINC* is expected to have a positive coefficient because students from wealthier districts have consistently been shown to perform better on standardized tests. The coefficient on *ABBOT* is expected to be negative because the criterion for inclusion in that category is *a priori* poor performance. Since higher ordinal rankings (i.e., 1–8) of *WDFG* capture more privileged students, *WDF*G is expected to be positively associated with test score performance. Finally, *GEOCEI* has an expected positive sign because it may proxy for family income differences in the different geographic areas of the state.

To aid in identification of the model, an additional uncontrollable socio-economic explanatory variable, student mobility, was added to the test outcome function. Student mobility (*STMOB*) is known to have a strong negative influence on test score outcomes (Hanushek et al. 2004).

The set of controllable management control tools in the outcome equation is the same as was used in the cost equation. Although the evidence on the pupil-teacher ratio is mixed (Hanushek 1986), we expect the coefficient on *STUFAC* to be negative according to conventional wisdom that smaller classes will enhance learning. The variable *STUCOMP* is also expected to be negative under the assumption that greater penetration of computers in the classroom enhances learning. The attendance rate (*ATTD*) is predicted to be positively related to test scores.

The three faculty-related variables, *FACSAL*, *CSINS*, and *FACEXP*, all have predicted positive signs in the test outcome equation under the expectation that higher-paid teachers, more resources devoted to teaching (relative to other functional areas), and more experienced teachers will all result in higher student achievement. For similar reasons, the coefficients on the three administrative-related factors, *ADMSAL*, *CSADM*, and *ADMEXPR* are also predicted to be positive.

The full test (quality) outcome function estimated can be written as:

$$\begin{aligned}\ln(CTEST) = b_0 &+ b_1 ELEM + b_2 MIDD + b_3 HIGH + b_4 \ln(EXPPP)\\ &+ b_5 \ln(WDFG) + b_6 \ln(GEOCEI) + b_7(\ln STMOB)\\ &+ b_8 \ln(SPED) + b_9 \ln(LEP) + b_{10} \ln(HIGHINC)\\ &+ b_{11} ABBOT + \gamma_1 \ln(STUFAC) + \gamma_2 \ln(STUCOMP)\\ &+ \gamma_3 \ln(ATTD) + \gamma_4 \ln(FACSAL) + \gamma_5 \ln(CSINS)\\ &+ \gamma_6 \ln(FACEXPR) + \gamma_7 \ln(ADMSAL)\\ &+ \gamma_8 \ln(CSADM) + \gamma_9 \ln(ADMEXPR) + e \qquad (6)\end{aligned}$$

When estimated together, the instrumental variables for *EXPPP* in Equation (5) are *ENROLL* and *SQ_ENROLL*, so these two variables do not appear in Equation (6). There is evidence in the literature to suggest that small schools are more cost-effective than larger schools (Toch 2003), but our preliminary regressions (consistent with results reported by other studies such as Childress et al. 2006) did not find such a relation. Conversely, the instrumental variable for *CTEST* is *STMOB*, which appears in Equation (6) but not in Equation (5).[18] There is no reason to believe that student mobility would have cost implications, other than student orientation costs, since no additional services need to be offered to students transferring into a school district or transferring out.

Given the apparent endogeneity of *EXPPP* and *CTEST*, we estimated several variations of Equations (5) and (6): as single equations individually using ordinary least-squares (OLS), jointly as a system of equations under seemingly unrelated regressions (SUR), and lastly as a system of simultaneous equations under three-stage least-squares (3SLS). We then performed the test of the fit of the instrumental variables, a statistical test for over-restriction of the instrumental variables, and the Hausman (1978) test for endogeneity.[19] We also tested for heteroscedasticity and multicollinearity.

RESULTS

Summary Statistics

Summary statistics on key variables used in the study are presented in Table 1 for year 2000. Because the profile of the school districts by type is similar in the other 2 years, they have been omitted for the sake of brevity.

Table 1 presents summary statistics on the distributions of per-pupil expenditures, total enrollments, and median faculty salaries, by type of school district. The following may be noted from these results. First, per-pupil expenditures vary systematically by type of school district. The lowest mean per-pupil expenditures occur in elementary school districts ($7,798), and increase progressively through middle school districts ($7,982), Kindergarten–Grade 12 districts ($8,512), and high-school districts ($9,918). Mean school district enrollments tend to be lowest for the *ELEM* and *MIDD* school districts, largest for the K–12 districts, with the *HIGH* school districts in between. Finally, median faculty salaries are highest for *HIGH*, followed by K–12, and lowest for the *MIDD* and *ELEM* school districts.

Table 2 presents the summary statistics of the variables used in this study pooled across school district types by year. Total student enrollments varied from 83 to 41,378 students in 2000, with similar ranges observed in 2001 and 2002. The proportion of special education needs students (*SPED*) ranged from 5 percent to a maximum of 28 percent in 2000, with the mean at 13 percent. By 2002, the maximum *SPED* had increased to 35 percent, although the mean was unchanged. Similarly, the proportion of limited English proficiency students (*LEP*) ranged from 0 to 27 percent in 2000 with the mean at 2 percent. The mean was virtually unchanged in 2001 and 2002, although the maximum increased to 39 percent in 2002.

[18] Note there are two instrumental variables for *EXPPP*, so Equation (6) is over-identified. There is one instrumental variable for *CTEST* so Equation (5) is exactly identified.

[19] Following Larcker and Rusticus (2005) and Hahn and Hausman (2002, 2003), we applied these three tests in succession. First, we checked for the weak instruments problem identified by Hahn and Hausman (2003), followed by the test for over-identifying restrictions for Equation (6). Finally, where the test for over-identifying restrictions indicated that the instruments were valid, we conducted a standard Hausman (1978) test to determine if a simultaneity problem existed.

TABLE 1
Summary Statistics of Expenditures, Enrollment, and Faculty Salaries by Type of School District Year 2000

	Number of Districts	Mean	Median	Minimum	Maximum
Expenditures Per Pupil by Type of District					
Grades Kindergarten to 6 (*ELEM*)	66	$7,798	$7,537	$5,214	$12,908
Grades Kindergarten to 8 (*MIDD*)	223	$7,982	$7,683	$5,355	$15,572
Grades Kindergarten to 12 (*K12*)	215	$8,512	$8,409	$6,156	$13,981
Grades 7/9 to 12 (*HIGH*)	47	$9,918	$9,821	$7,235	$13,959
Total Enrollments by Type of District					
Grades Kindergarten to 6 (*ELEM*)	66	588	359	60	3,687
Grades Kindergarten to 8 (*MIDD*)	223	894	615	84	7,927
Grades Kindergarten to 12 (*K12*)	215	4,313	3,030	530	41,378
Grades 7/9 to 12 (*HIGH*)	47	1,666	1,369	386	8,589
Median Faculty Salaries by Type of District					
Grades Kindergarten to 6 (*ELEM*)	66	$41,277	$42,151	$9,843	$55,860
Grades Kindergarten to 8 (*MIDD*)	223	$45,678	$44,930	$29,693	$68,377
Grades Kindergarten to 12 (*K12*)	215	$52,086	$50,914	$35,780	$69,450
Grades 7/9 to 12 (*HIGH*)	47	$55,999	$55,337	$41,687	$80,710

Among the controllable management tools, the student-faculty ratio averaged 12.69 in 2000, 12.48 in 2001, and 12.28 in 2002, a steady decline that is also noticeable with students-computers ratio (6.35, 5.89, and 4.98 in the three years, respectively). However, the attendance rate held steady at an average of 95 percent in all the three years. The cost shares for instruction and administration (*CSINS* and *CSADM*, respectively) also remained constant, with averages of 61 percent and 13 percent respectively for all the three years. The level of faculty experience (*FACEXPR*) experienced a decline from a mean of 14.67 years in 2000, to 13.97 years in 2001, and finally to 13.17 years. This slight decline is also noticeable for administrative experience (*ADMEXPR*), which averaged 23.91, 23.86, and 23.53 for the three years, respectively, although the medians were the same. These results may reflect the result of personnel attrition as well as intra-state and inter-state mobility.

As noted previously, the ability of school district administrators to initiate changes in, for example, the average teacher experience may be limited by teacher-union contracts, tenure rules, and so forth. Nevertheless, flexibility exists for district administrators to enforce institute changes by hiring only inexperienced teachers, shifting resources toward instructional costs or other functional areas, and enacting policies to increase student attendance, for example. The effectiveness of such policies relative to the cost of enforcement is one focus of this study.

Results from Estimating the Cost Function

The results of estimating the cost function specified in Equation (5) using the SUR approach are presented in Table 3. Because the results obtained using OLS, 2SLS, and

TABLE 2
Summary Statistics of Variables used in the Study Year 2000

Variable	Abbreviation	Mean	Median	Minimum	Maximum
Endogenous Variables					
Total Expenditures	*TCOST*	$20,002,157	$10,617,203	$683,556	$453,257,853
Expenditures Per Pupil	*EXPPP*	$8,363	$8,129	$5,214	$15,034
Composite Average Test Score	*CTEST*	0.88	0.88	0.71	1.00
Uncontrollable Variables					
Total Student Enrollments	*ENROLL*	2,346	1,234	83	41,378
Square of Log of Student Enrollments	*SQ_ENROLL*	52.17	50.67	19.53	113.01
Weighted District Factor Scores	*WDFG*	4.31	4.00	0.69	8.00
Geographic Cost of Education Index	*GEOCEI*	1.01	1.00	0.83	1.13
Student Mobility Rate	*STMOB*	11.20	9.00	0.80	57.80
Percent Students in Special Education Programs	*SPED*	0.13	0.12	0.05	0.28
Percent Students in English Proficiency Programs	*LEP*	0.02	0.01	0.00	0.27
Percent Students not Receiving Federal Aid	*HIGHINC*	0.86	0.93	0.15	1.00
Abbot District Membership	*ABBOT*	0.05	0.00	0.00	1.00
Controllable Variables					
Student Faculty Ratio	*STUFAC*	12.69	12.60	6.05	20.50
Student-Computer Ratio	*STUCOMP*	6.35	5.40	1.20	73.80
Student Attendance Rate	*ATTD*	0.95	0.95	0.88	0.98
Average Faculty Salary	*FACSAL*	$49,080	$47,731	$29,179	$80,710
Cost Share of Instructional Expenses	*CSINST*	0.61	0.61	0.47	0.73
Average Faculty Experience	*FACEXPR*	14.67	15.00	5.00	26.00
Average Administrative Salary	*ADMSAL*	$80,235	$81,127	$44,427	$109,650
Cost Share of Administrative Expenses	*CSADM*	0.13	0.12	0.06	0.26
Average Administrative Experience	*ADMEXPR*	23.91	25.00	2.00	39.00

3SLS are similar to the SUR results, we omit them for the sake of brevity.[20] For ease of referencing, the natural log transformation of the variables will be ignored when referring to these variables in the regression. Thus, ln(*EXPPP*), for example, will be referred to simply as *EXPPP*.

Panels A, B, and C of Table 3 present the results for years 2000, 2001, and 2002, respectively. The reported significance levels are one-tailed for variables where a specific sign was expected, and two-tailed where there was no *a priori* expectation. The results in all three panels show that, with the exception of the limited English proficiency ratio

[20] Seemingly unrelated regression (SUR) is a full information statistical approach that exploits the likely cross-equation correlation of the error terms. For well-specified models, SUR is more efficient than OLS (Greene 1990). Its simultaneous equation counterpart is three-stage least-squares (3SLS), which applies the SUR approach to the 2SLS estimates. We applied all four possible statistical methods, and have chosen to focus on the SUR results because the coefficient estimates are more efficient and less bias.

TABLE 3
Results of Regression of Per-Pupil Expenditure (*ln_EXPPP*) on Determinants of School District Spending using Seemingly Unrelated Regression (SUR)[a]

		Panel A: Year 2000		Panel B: Year 2001		Panel C: Year 2002		Variable Inflation Factors
Variable[b]	**Expected Sign[c]**	**Coefficient**	**t-value Significant Level**	**Coefficient**	**t-value Significant Level**	**Coefficient**	**t-value Significant Level**	**Year 2000**
Intercept		4.947	8.90***	5.480	10.37***	7.579	14.85***	0
ELEM (K–6)	–	−0.080	−3.20***	−0.093	−3.73***	−0.040	−1.74***	3.77
MIDD (K–8)	–	−0.100	−7.34***	−0.092	−7.12***	−0.083	−6.56***	2.79
HIGH (7–12 and 9–12)	+	0.024	1.29&	0.009	0.51	0.022	1.33&	1.80
Jointly Endogenous								
ln_CTEST	+	0.390	2.8***	0.818	5.75***	0.613	3.97***	3.78
Uncontrollable Variable								
ln_ENROLL	–	−0.372	−8.35***	−0.330	−7.46***	−0.350	−8.16***	157.26
SQ_ENROLL	+	0.020	6.78***	0.017	5.84***	0.019	6.92***	145.06
ln_WDFG	+	0.039	2.52**	0.042	2.76***	0.067	4.63***	6.29
ln_GEOCEI	+	0.413	3.69***	0.326	3.17***	0.567	5.67***	3.43
ln_SPED	+	0.122	6.62***	0.072	4.23***	0.067	4.03***	1.47
ln_LEP	+	0.003	0.65	0.004	0.97	0.004	1.21	2.34
ln_HIGHINC	?	−0.107	−3.40***	−0.164	−4.97**	−0.164	−4.93***	4.49
ABBOT	–	0.139	5.43***	0.081	3.32***	0.164	6.96***	2.11

(*continued on next page*)

TABLE 3 (continued)

Variable[b]	Expected Sign[c]	Panel A: Year 2000 Coefficient	Panel A: Year 2000 t-value Significant Level	Panel B: Year 2001 Coefficient	Panel B: Year 2001 t-value Significant Level	Panel C: Year 2002 Coefficient	Panel C: Year 2002 t-value Significant Level	Variable Inflation Factors Year 2000
Controllable Variable								
ln_STUFAC	−	−0.308	−10.81***	−0.308	−11.67***	−0.363	−14.00***	1.76
ln_STUCOMP	−	−0.028	−2.76***	−0.024	−2.56**	−0.055	−5.19***	1.53
ln_ATTD	?	0.318	0.70	−0.342	−0.77	−0.166	−0.39	2.33
ln_FACSAL	+	0.375	8.95***	0.371	9.47***	0.324	8.95***	3.41
ln_CSINS	?	−0.664	−7.30***	−0.764	−9.16***	−0.523	−6.60***	2.10
ln_FACEXPR	+	0.001	0.03	−0.005	−0.27	−0.002	−0.10	2.18
ln_ADMSAL	+	0.206	4.94***	0.125	3.09***	0.014	0.37	2.86
ln_CSADM	?	−0.086	−2.80***	−0.135	−4.79***	−0.096	−3.55***	2.54
ln_ADMEXPR	?	−0.020	−1.09	−0.001	−0.05	0.006	0.47	1.55
Number of Observations			520		520		521	
Adjusted R^2 (OLS)			0.721		0.710		0.744	

, *, & Significant at 0.01 percent, 0.001 percent, and 0.10 percent levels, respectively.
[a] SUR results of Equation (5) estimated jointly with Equation (6) presented in Table 4.
[b] All variables are defined in Exhibit 1.
[c] Hypotheses tests on coefficients are one-tailed unless the expected sign of coefficient is undeterminable (then two-tailed tests are used).

EXHIBIT 1
List of Variables Used in the Study

Variable	Definition
District Type	
K12	Dummy variable that equals 1 for Kindergarten to Grade 12 School Districts.
ELEM	Dummy variable that equals 1 for Elementary School Districts (Grades Kindergarten to 6).
MIDD	Dummy variable that equals 1 for Middle School Districts (Grades Kindergarten to 8).
HIGH	Dummy variable that equals 1 for High School Districts (Grades 7–12 and Grades 9–12).
Endogenous	
ln_TCOST	Total operating expenditures for the school district.
ln_EXPPP	Expenditures per pupil, defined as total operating expenditures divided by average daily enrollment.
ln_CTEST	Composite weighted average test scores, defined in Equation (3).
Exogenous Uncontrollable	
ln_ENROLL	Total school district student enrollment.
SQ_ENROLL	Square of *ln_ENROLL*.
ln_WDFG	Weighted school district factor grouping index.
ln_GEOCEI	Geographical cost of education index.
ln_STMOB	Student mobility ratio.
ln_SPED	Ratio of Special Education students to total student enrollment.
ln_LEP	Ratio of Limited English Proficiency students to total student enrollment.
ln_HIGHINC	Complement of the percentage of students in the school district receiving meal aid under the federally subsidized school lunch program.
ABBOT	Dummy variable that equals 1 for Abbott school districts.
Controllable	
ln_STUFAC	Ratio of number of students to number of faculty in the school district.
ln_STUCOMP	Ratio of number of students to number of computers in the school district.
ln_ATTD	Average class attendance rate for the school district.
ln_FACSAL	Average faculty salary in the school district.
ln_CSINS	Cost share for instruction, computed as total instructional expenditures divided by total operating expenditures.
ln_FACEXPR	Average number of years of teacher experience in the school district.
ln_ADMSAL	Average salary of administrators in the school district.
ln_CSADM	Cost share for administration, computed as total administrative expenditures divided by total operating expenditures.
ln_ADMEXPR	Average number of years of administrator's experience in the school district.

(LEP), all the uncontrollable variables are significant and consistent with *a priori* expectations, where such expectations existed. Since the results are consistent across all three years, we will use the year 2000 results to illustrate the main findings in this section. The negative coefficient for *ENROLL* (student enrollments) is consistent with the expectation of increasing economies of scale (i.e., $\alpha_5 = 0.628$ since $1 - \alpha_5 = -0.372$ in 2000). At the same time, the positive coefficient for *SQ_ENROLL* (although with a smaller coefficient of

0.02) reflects diseconomies of scale for the largest school districts. The negative coefficient for *HIGHINC* (complement of percent students receiving meal aid) is consistent with the interpretation that school districts in high-income brackets tend to be more cost-efficient than districts in poor income areas when other factors are controlled for.

Of the nine management control variables, all but three are significant across all three years. The three exceptions are the attendance rate (*ATTD*), faculty experience (*FACEXPR*), and administrator experience (*ADMEXPR*). Of the six controllable variables with significant signs, *STUFAC* (average class size) and *STUCOMP* (number students per computer) are negative as expected. Similarly, average faculty salary (*FACSAL*) and average administrator's salary (*ADMSAL*) are both positively signed as expected. Of the remaining two variables whose sign could not be established on an *a priori* basis, both the cost shares for instruction and administration (*CSINS* and *CSADM*, respectively) have negative coefficients. This implies that cost efficiency is associated with higher spending on both instruction and administration at the expense of other costs (such as operating and maintenance, and student support costs). The intriguing question is whether higher levels of these two inputs are related to achievement (quality) outcomes. This issue is addressed in the next section.

The final column of Table 3 presents the variance inflation factors (VIFs) obtained in the cost equation regression for 2000. The VIFs gauge the degree of multicollinearity in the regression estimates, and are defined as the inverse of the multiple correlation coefficients from the regression of all the independent variables on all the other independent variables (Belsley et al. 1980). The VIFs for years 2001 and 2002 are similar and have been omitted. Results indicate that except for the high degree of correlation between *ENROLL* and *SQ_ENROLL*, none of the other variance inflation factors are a cause for concern.

In order to investigate the possible simultaneous equation bias issue, Equation (5) was estimated jointly with Equation (6) using the 3SLS approach. The results using 3SLS (not reported) are substantially identical to those presented in Table 3 based on SUR. The main difference is the coefficient on weighted-average test scores (*CTEST*). This variable was significant in the SUR model in Table 3 ($p < 0.001$), but is not significant under 3SLS.

Specification tests to determine if simultaneous equation bias existed were performed on the 3SLS results. The test for the fit of the instrumental variables is based on partial F-statistics, which measures the incremental contribution of the instruments in the first-stage regressions. The partial F-statistics are 37.27, 27.54, and 19.45 for years 2000, 2001, and 2002, respectively. Thus, for all three years, the instrument used for *CTEST* (namely *STMOB*—student mobility) was found to be strong. Since Equation (5) was exactly identified, there was no need to perform the test for over-identified restrictions. Proceeding to the Hausman (1978) test shows that the null hypothesis of no endogeneity cannot be rejected in any of the three years, with F ratios of 0.52, 0.06, and 0.00, respectively. These results indicate that the SUR results are less biased than the 3SLS results for Equation (5). Since the SUR results in Table 3 shows *CTEST* to be statistically significant while the 3SLS results did not, the results support the conclusion that test scores have an independent effect on spending levels.

Results from Estimating the Test Outcome Function

Results of estimating the outcome function in Equation (6) under SUR are presented in Table 4, with years 2000, 2001, and 2002 in Panels A, B, and C, respectively.

All three panels of Table 4 indicate that three of the uncontrollable variables (*WDFG*, *STMOB*, and *HIGHINC*) are significant in all years. *WDFG* (weighted district factor group

TABLE 4

Results of Regression of Test Scores (*ln_CTEST*) on Determinants of Student Performance on using Seemingly Unrelated Regression (SUR)[a]

		Panel A: Year 2000		Panel B: Year 2001		Panel C: Year 2002		Variable Inflation Factors
Variable[b]	**Expected Sign[c]**	**Coefficient**	**t-value Significant Level**	**Coefficient**	**t-value Significant Level**	**Coefficient**	**t-value Significant Level**	**Year 2000**
Intercept		−0.522	−2.93***	−0.937	−5.72***	−0.976	−6.08***	0.00
ELEM (K–6)	?	0.042	5.87***	0.058	8.58***	0.035	5.63***	3.27
MIDD (K–8)	?	0.002	0.59	0.014	3.80***	0.009	2.83***	2.33
HIGH (7–12 and 9–12)	?	0.007	1.23	−0.001	−0.20	0.001	0.13	1.76
Jointly Endogenous								
ln_EXPPP	+	0.036	3.01***	0.065	5.58***	0.044	3.93*	2.84
Uncontrollable Variable								
ln_WDFG	+	0.041	9.17***	0.034	7.77***	0.027	6.88***	5.43
ln_GEOCEI	?	−0.072	−2.05*	−0.055	−1.76	−0.041	−1.42	3.50
ln_STMOB	−	−0.017	−6.05***	−0.011	−4.89***	−0.009	−4.26***	2.23
ln_SPED	−	−0.010	−1.61&	−0.003	−0.64	−0.009	−1.87*	1.57
ln_LEP	?	0.003	2.36*	0.002	1.48	0.001	1.00	2.55
ln_HIGHINC	+	0.044	4.61***	0.060	6.18***	0.062	6.91***	4.24
ABBOT	−	−0.022	−2.76**	−0.017	−2.35**	−0.005	−0.69	2.10

(*continued on next page*)

TABLE 4 (continued)

Variable[b]	Expected Sign[c]	Panel A: Year 2000 Coefficient	Panel A: Year 2000 t-value Significant Level	Panel B: Year 2001 Coefficient	Panel B: Year 2001 t-value Significant Level	Panel C: Year 2002 Coefficient	Panel C: Year 2002 t-value Significant Level	Variable Inflation Factors Year 2000
Controllable Variable								
ln_STUFAC	−	0.005	0.46	0.024	2.69***	0.023	2.66**	2.17
ln_STUCOMP	−	−0.008	−2.58**	−0.005	−1.76*	−0.003	−0.89	1.51
ln_ATTD	+	0.395	2.80***	0.387	2.92***	0.478	4.05***	2.30
ln_FACSAL	+	−0.008	−0.58	−0.003	−0.21	0.002	0.16	3.83
ln_CSINS	+	0.084	2.95***	0.086	3.28***	0.014	0.63	2.15
ln_FACEXPR	+	−0.001	−0.18	−0.008	−1.35	−0.005	−1.09	2.17
ln_ADMSAL	+	0.021	1.65&	0.027	2.24**	0.038	3.70***	2.78
ln_CSADM	+	0.004	0.43	0.001	0.07	−0.001	−0.16	2.08
ln_ADMEXPR	+	−0.011	−1.99*	−0.009	−2.03*	−0.008	−2.22*	1.54
Number Observations			520		520		521	
Adjusted R^2 (OLS)			0.745		0.761		0.749	

*, **, ***, & Significant at 0.05 percent, 0.01 percent, 0.001 percent, and 0.10 percent levels, respectively.
[a] SUR results of Equation (6) estimated jointly with Equation (5) presented in Table 3.
[b] All variables are defined in Exhibit 1.
[c] Hypotheses tests on coefficients are one-tailed unless the expected sign of coefficient is undeterminable (then two-tailed tests are used).

index) and *HIGHINC* have the expected positive sign, while *STMOB* is negative (as expected). Of the remaining four uncontrollable variables, the dummy variable for Abbott districts (*ABBOT*), the geographical cost of education index (*GEOCEI*), and the ratio of special education students to the student population (*SPED*) are all negatively signed and significant in at least one of the three years. *LEP* is positively signed but significant in only year 2000.

Of the nine management control variables, three are consistently significant across all three years: *ATTD* and *ADMSAL* (with expected positive coefficients), and *ADMEXPR* (with an unexpected negative coefficient). Of the remaining six controllable variables, *STUFAC*, *STUCOMP*, and *CSINS* are significant in 2 of the three years. In contrast to expectations, *STUFAC* has a significant positive coefficient in 2001 and 2002, suggesting that students perform better in larger classes. *FACSAL*, *CSADM*, and *FACEXPR* have statistically insignificant coefficients in all three years. Finally, the jointly endogenous variable, expenditures per-pupil (*EXPPP*), is strongly significant in all three years.

The final column of Table 4 presents the variance inflation factors (based on the 2000 data) to permit the degree of multicollinearity to be assessed. Since none of the observed values exceed six (the threshold provided by Belsley et al. (1980)), and the results for 2001 and 2002 are similar to those reported here, multicollinearity is not a problem in these regressions.

Like the cost function, tests for specification of the outcome function were conducted by estimating Equation (6) simultaneously with Equation (5) using 3SLS. The partial F-statistics for evaluating the incremental contribution of the instrumental variables are 65.21, 80.31, and 55.49 for 2000, 2001, and 2002, respectively. The instruments have high explanatory power, alleviating any concern about potentially weak instruments. In addition, the χ^2 values from the over-identifying restrictions test were 0.21, 0.05, and 0.36 for the three years respectively, none of which are significant at reasonable probability levels. Thus, the instruments can be regarded as exogenous. Finally, the F-ratios from the Hausman (1978) test were 0.03, 0.31, and 0.010, respectively, all of which are statistically insignificant. The null hypothesis of no endogeneity cannot be rejected. In conclusion, the SUR results are less biased than the 3SLS results, consistent with the earlier findings. For this reason, presentation of the 3SLS results are omitted from this paper.[21]

Summary and Synthesis of Results

Before proceeding to an analysis of the results, it is important to note that the least-squares regressions used here, unlike the stochastic (and/or deterministic) frontier estimation techniques used by D&G (1997) and Mensah and Li (1993), among others, are not based on extreme values. That is, the least-squares method used in this paper estimates the average effects of unit changes in the independent variables on the dependent variable within a relevant range. This choice is deliberate since we seek to provide insight into the trade-off decisions made by the average school superintendent who is facing the multiple constituencies and objectives. We do not make the assumption that the average school superintendent is a cost-minimizer or outcome-maximizer.

To synthesize the results provided earlier, the relative cost-effectiveness indicators (coefficients) for the controllable variables estimated under SUR are summarized in Table 5

[21] In contrast to the strong results observed for *EXPPP* in Table 4, the coefficient on *EXPPP* from the 3SLS regression is not significant. However, the statistical insignificance of the Hausman test for endogeneity suggests that these results can be rejected in favor of the SUR results in Table 4.

TABLE 5
Analysis of the Cost Effectiveness of Alternative Controllable Measures in Public School Administration SUR Results

Panel A: 2000

Variable[b]	Effect on Costs Significant Level	Effect on Test Scores Significant Level	Relative Cost Effectiveness[a]	Trade-Off Ratio
ln_STUFAC	−0.308***	0.005	BC	
ln_STUCOMP	−0.028***	−0.008**	TO (NEG)	0.286
ln_ATTD	0.318	0.395***	BO	
ln_FACSAL	0.375***	−0.008	CI	
ln_CSINS	−0.664***	0.084***	DB	
ln_FACEXPR	0.001	−0.001	NS	
ln_ADMSAL	0.206***	0.021&	TO (POS)	0.102
ln_CSADM	−0.086***	0.004	BC	
ln_ADMEXPR	−0.020	−0.011*	NO	

Panel B: 2001

Variable[b]	Effect on Costs Significant Level	Effect on Test Scores Significant Level	Relative Cost Effectiveness[a]	Trade-Off Ratio
ln_STUFAC	−0.308***	0.024***	DB	
ln_STUCOMP	−0.024**	−0.005*	TO (NEG)	0.208
ln_ATTD	−0.342	0.387***	BO	
ln_FACSAL	0.371***	−0.003	CI	
ln_CSINS	−0.764***	0.086***	DB	
ln_FACEXPR	−0.005	−0.008	NS	
ln_ADMSAL	0.125***	0.027**	TO (POS)	0.216
ln_CSADM	−0.135***	0.001	BC	
ln_ADMEXPR	−0.001	−0.009*	NO	

Panel C: 2002

Variable[b]	Effect on Costs Significant Level	Effect on Test Scores Significant Level	Relative Cost Effectiveness[a]	Trade-Off Ratio
ln_STUFAC	−0.363***	0.023**	DB	
ln_STUCOMP	−0.055***	−0.003	BC	
ln_ATTD	−0.166	0.478***	BO	
ln_FACSAL	0.324***	0.002	CI	
ln_CSINS	−0.523***	0.014	BC	
ln_FACEXPR	−0.002	−0.005	NS	
ln_ADMSAL	0.014	0.038***	BO	
ln_CSADM	−0.096***	−0.001	BC	
ln_ADMEXPR	0.006	−0.008*	NO	

*, **, ***, & Significant at 0.05 percent, 0.01 percent, 0.001 percent, and 0.10 percent levels, respectively.
[a] Relative Cost Effectiveness Key:
BC = beneficial to costs, but no effect on test score outcomes;
BO = beneficial to test score outcomes, but no effect on costs;
DB = doubly beneficial (i.e., beneficial effects on both costs and test score outcomes);

(*continued on next page*)

TABLE 5 (continued)

DN = doubly negative (i.e., negative effect on both costs and test score outcomes);
CI = cost increasing, with no off-setting beneficial effect on outcomes;
NO = negative effect on test score outcomes, but no effect on costs;
TO = trade-off effect (i.e., either a beneficial effect on costs offset by negative effect on test score outcomes, or vice versa), figure in parenthesis is: Relative Rate of Trade-Off of Outcomes per Cost = +/− |Outcome coefficient|/|Cost coefficient|; and
NS = no statistically significant effect on costs or outcomes.
[b] All variables are defined in Exhibit 1.

for each of the three years. For the cost function results, positive coefficients indicate cost-increasing factors, while negative coefficients are cost-beneficial. In contrast, for the test outcome function, positive coefficients are outcome beneficial, while negative signs have the opposite effect.

The results in Table 5 show that, of the various factors that are controllable by school administrators, four deserve special favorable attention. First, *CSINS* is doubly beneficial (DB = outcome-increasing and cost-decreasing at the same time) in 2000 and 2001, and rates a BC (beneficial cost effect) in 2002. This is closely followed by *ATTD*, which has a BO (beneficial effect on outcomes with no offsetting cost-increasing effect) in all three years. Somewhat surprisingly, *STUFAC* earns a BC (cost-decreasing effect with no off-setting negative effect on outcomes) rating in 2000, and a DB rating in 2001 and 2002. However, we should note that the state of New Jersey has a mandatory maximum student/faculty ratio for the different types of school districts and grades. Subject to this state mandate, it appears that, for the data and the years covered, school districts whose student/faculty ratios were above the state-wide means (but which met the state-wide maximum limits) actually did better from both a cost-efficiency and test outcome standpoint than schools with lower student/faculty ratios.[22] This finding suggests that school districts significantly below the state-mandated maximum student/faculty ratios did not benefit much from those decisions.

The fourth variable of note is *CSADM*, which earned a BC (cost-reducing effect with no negative effect on outcomes) rating in all three years. Once again, it is important to note that the New Jersey Department of Education monitors the administrative cost ratios and has guidelines regarding them. Granted this context, the results suggest that schools whose administrative cost ratios were above the state-wide means did, in fact, achieve better control over costs without paying a penalty in terms of lower test score performance. A closer examination of this finding and its implications clearly merit attention in future research.[23]

At the opposite end of the spectrum, two variables stand out as highly negative in their impact on either costs or outcomes. *ADMEXPR* has a uniformly NO interpretation (negative outcome effect with no offsetting effect on costs) in all years. This result suggests that, at

[22] The case for *STUFAC* is more complicated than in the results reported here because the data is pooled across the four different types of school districts. In a more detailed analysis based on disaggregated data, we found *STUFAC* in the *K12* and *ELEM* school systems to be a TO with a negatively signed relative cost-effectiveness measure. However, the trade-off ratios ($|\gamma_i|/|\lambda_i|$) were quite small, indicating that the dampening effect of higher student/faculty ratios on outcomes is modest compared to their cost savings effect. (That is, they contribute more to cost savings than they induce lower test-score performance.) In the *HIGH* and *MIDD* school districts, *STUFAC* earned a straight BC (cost-savings with no off-setting negative effect on outcomes). This suggests that relatively high levels of *STUFAC* may actually be more cost-effective than excessively low ratios.

[23] It should be noted that, in addition to institutional constraints placed on the cost share for administration by the state of New Jersey, taxpayer preferences may influence other outcomes not captured by our model. For example, the desire to produce high-quality athletic teams may effectively limit the ability of school district administrators to reallocate funds to administration and instruction and away from extra-curricular activities.

the margin, an increase in *ADMEXPR* is associated with a decrease in test scores, but simultaneously has a positive effect on costs. Thus, one inference from these results is that schools with *ADMEXPR* above the state-wide mean may benefit from the flexibility to hire administrators with less experience who command lower salaries, with all other factors held constant. The retirement of older, more experienced administrators, if they are replaced by less experienced but lower-paid administrators, may actually be beneficial to such school districts.

The other noticeable negative factor is *FACSAL*, which is rated CI (cost increasing with no offsetting benefit to outcomes) in all three years. These results suggest that school districts with faculty salaries above the state-wide means do not obtain any extra payoff from the higher salaries. One intuitive interpretation for this finding is that collective bargaining agreements with the teachers' union may limit the ability of administrators to control teacher salaries.

Finally, two of the nine management tools have ratings of TO (trade-off of costs against outcomes) in the first 2 years, with a change to either BC or BO in 2002. The first variable, *STUCOMP*, is the number of students per computer. *STUCOMP* has a negative TO rating in 2000 and 2001, implying that greater utilization of computers in the classroom increases total costs, but also increases test scores. The magnitude of the trade-off for *STUCOMP* decreased from 0.286 in 2000, to 0.208 in 2001, but in 2002 *STUCOMP* had a rating of BC (beneficial to costs with no offsetting reduction of outcomes). As noted earlier, the student/computer ratio exhibited a steady decrease over the sample period, reflecting a trend toward the introduction of more computers in the classroom. The change in the rating for *STUCOMP* may be related to this decrease, but additional years of data and/or further investigation would be required to establish any causal relation.

The other variable with a TO rating is *ADMSAL*. The positive TO ratings in 2000 and 2001 indicate that the increased effect on outcomes due to greater administrative salaries is offset by its cost-increasing effect. The TO rating for *ADMSAL* increases from 0.102 in 2000 to 0.216 in 2001, and is finally BO (beneficial on outcomes only) in 2002 under the SUR results. This was accomplished through a steady increase in the outcome coefficient, and a similarly steady decrease in the cost coefficient. However, since one cannot rule out the possibility that the trends in the ratios over time are merely statistical noise, no further interpretation can be offered without additional years of data.

Contribution of Joint Estimation Methodology

As noted in the introduction, the aim of this paper is to propose the joint estimation of cost and outcome functions to enable management accountants to understand the coping mechanisms adopted by non-profit managers who face multiple constituents and multiple objectives, some of which may be in conflict. The results show that, for controllable variables denoted BC, BO, or DB, decision unit managers appear to be utilizing low levels of beneficial resources, within the range of our data set. In still other cases, some managers utilize high levels of resources identified as NO or CI that on average provide no further benefit within the range. Clearly, one possible explanation for such behavior might be if the decision unit managers face institutional constraints that limit their freedom of action in important ways.

The proposed joint estimation methodology thus allows specific resources or control tools whose levels appear to be under or over utilized, within the relevant range of our data set, to be identified. Further investigation by the management accountants would permit the exact nature of the institutional constraints (and the associated opportunity costs) to be identified and possibly dealt with by the decision unit manager. Since the primary contribution of the paper is intended to be methodological, and the school district example was only used for illustrative purposes, we provide below only a brief overview of how the method can be beneficial in this context.

The literature on public school education is vast and increasing in scope, particularly following national education initiatives such as the No Child Left Behind Act of 2001 (U.S. Department of Education 2002). However, what is most striking about this vast literature is the consistency with which contradictory results have been found. In a thorough review of the education literature, Hanushek (1986) observed the following:

(1) Out of 112 studies of outcome functions surveyed, 23 studies found the Teacher/Pupil ratio to be a statistically significant variable, while 89 found it to be insignificant. Of the 23 that reported statistical significance, nine reported positive coefficients, while 14 reported negative coefficients. (We found *STUFAC* to be positive and significant in our outcome function in 2001 and 2002, but not significant in 2000).[24]
(2) In the 109 studies where *FACEXPR* was examined in the outcome function, 40 studies found it to be statistically significant, while 69 found it to be insignificant. Of these 40 studies, 33 reported a positive coefficient, while 7 reported a negative coefficient. (We found *FACEXPR* to be statistically insignificant across all years in the outcome function).
(3) In 60 studies where *FACSAL* was included as an independent variables in the outcome function, 10 reported statistically significant coefficients (nine positive signs, and one negative sign) while 50 reported no statistically significant coefficients. (We found the variable *FACSAL* had no effect on test scores in all three years).
(4) In 65 studies that included per-pupil expenditures in the outcome function, 16 studies found it to be a significant explanatory variable. Of these 16, 13 studies had positive coefficients and three had negative coefficients. (We found *EXPPP* to be significant and positive across all three years in our outcome function under SUR estimation).

These studies summarized by Hanushek (1986) span a number of decades, and reflect institutional factors specific to different states and even evolving national policies. As a result, the findings of educational studies may not be generalizable beyond the time periods and samples from which the data are obtained.[25] Additionally, the studies may use different measures for explanatory variables and employ different statistical models. Educational outcomes are also often based on different achievement tests. These differences undermine any possibility that this study (or any other, for that matter) can yield authoritative findings that can resolve the overwhelming conflicts reported in the literature.

[24] The variable used in our models, *STUFAC*, is the inverse of the Teacher/Pupil ratio. Thus, our finding of a positive coefficient on *STUFAC* in years 2001 and 2002 in our outcome function is consistent with the 14 studies summarized by Hanushek (1986) that found a negative coefficient on the Teacher/Pupil ratio.

[25] Our assertion that the results of educational studies may not be replicable outside of the sample periods and locations is supported by the fact that when we attempted to replicate D&G (1997) using data from the state of New Jersey (instead of the state of Missouri data), the coefficients we derived (as well as the level of cost inefficiency) were very different from the ones they found.

CHECKS FOR ROBUSTNESS OF RESULTS

In this section, we perform checks to determine the robustness of our results. First, given the correlation between some of the independent variables, we examine how the omission of any of the variables affects the explanatory power of the remaining variables in the cost and outcome equations. Second, we evaluate more carefully the frequently held view that devoting more resources to instruction is the best way to make public education more cost-effective. Finally, we test the models for the underlying assumptions of linearity (in log terms) and additivity of the variables. All robustness tests are conducted by pooling the data cross-sectionally over years 2000–2002.

Incremental Explanatory Power

The coefficients on all the variables are essentially marginal effects obtained with all the other variables held constant at their mean values (Johnston 1972, 132–135). To aid in understanding the relative importance of the individual variables, we estimated the SUR regressions with one variable omitted at a time. The R^2s without the omitted variable and the incremental change in the OLS and system-weighted R^2s are disclosed in Table 6.

The results in Table 6 are sorted by the magnitude of the overall effect on the SUR system-weighted R^2. But for analytical purposes, the incremental effect of the omission of each variable on the individual OLS R^2s are more relevant. The results for the cost equation show that the controllable variables with the greatest incremental contribution in explaining total expenditures are the student-faculty ratio (7.5 percent), mean faculty salaries (4.6 percent), and total cost share of instruction (3.2 percent). None of the other controllable variables exceeds 1.0 percent. Of the noncontrollable variables, the economies of scale effect (*ENROLL* & *SQ_ENROLL*), had the most impact on the observed explanatory power of the model (7.4 percent).

Additional insights are provided in Table 6 by examining the most important individual factors that contribute to explaining test score performance. The most important variables in the outcome equation are the district factor groupings (*WDFG*, 0.033), level of district income (0.017), student mobility (0.017), and the existence of special education programs (0.017). None of the controllable variables have an incremental explanatory contribution to the model that exceeds 1.0 percent. This finding is consistent with the literature summarized by Hanushek (1986), which shows that once socio-economic factors are controlled for, school inputs such as the student-faculty ratio, teacher salaries, and per-pupil expenditures explain a relatively small portion of student achievement scores.

Linearity

The results presented thus far do not allow for possible non-linearity or curvatures in the model. However, it can be argued that for many of these variables, it is likely that increases in the quantities may push the relationship identified beyond levels that are beneficial. One good example is the cost share assigned to instruction. Although our results, along with the prior literature, accords with the view that as much resources as possible should be devoted to instruction, it is possible to envision situations where too many resources are devoted to instruction, and insufficient resources are made available to other functional areas. To examine this issue in more detail, we divided the pooled sample into three equal groups as follows: (1) *CSINS_1* denotes those observations with instructional cost share ranging from the minimum of 47.04 percent to 59.2 percent; (2) *CSINS_2* denotes observations with *CSINS* ranging from 59.2 percent to 62.2 percent; and (3) *CSINS_3* denotes those observations with instructional cost shares ranging from 62.2 percent to the maximum of 73.2 percent. Equations (5) and (6) were re-estimated with the three partitioned

TABLE 6
Effect on OLS and SUR System-Weighted R^2 of Omitting Dependent Variables
Pooled Cross-Sectional Results 2000–2002

	R^2 with Variable Omitted			Incremental R^2 Attributable to Omitted Variable		
Variables in Equations	**Cost Equation OLS**	**Test Score Equation OLS**	**SUR System Weighted**	**Cost Equation OLS**	**Test Score Equation OLS**	**SUR System Weighted**
R^2 including all variables in cost and test score equations[a]	0.724	0.746	0.745			
ln_STUFAC	0.649	0.745	0.713	0.075	0.001	0.032
ln_ENROLL and *SQ_ENROLL*	0.650	NA[b]	0.720	0.074	NA	0.025
ln_FACSAL	0.678	0.746	0.729	0.046	0.000	0.016
ln_CSINS	0.692	0.744	0.729	0.032	0.002	0.016
ln_HIGHINC	0.716	0.729	0.729	0.008	0.017	0.016
ln_WDFG	0.717	0.713	0.732	0.007	0.033	0.013
ln_CTEST	0.718	NA	0.733	0.006	NA	0.012
ln_EXPPP	NA	0.740	0.733	NA	0.006	0.012
ABBOTT	0.711	0.744	0.738	0.013	0.002	0.007
ln_STMOB	NA	0.729	0.738	NA	0.017	0.007
ln_SPED	0.724	0.729	0.738	0.000	0.017	0.007
ln_ADMSAL	0.716	0.741	0.740	0.008	0.005	0.005
ln_GEOCIE	0.717	0.744	0.740	0.007	0.002	0.005
ln_CSADM	0.717	0.746	0.742	0.007	NA	0.003
ln_ATTD	0.724	0.741	0.743	0.000	0.005	0.002
ln_STUCOMP	0.717	0.743	0.744	0.007	0.003	0.001
ln_LEP	0.724	0.744	0.744	0.000	0.002	0.001
ln_FACEXPR	0.724	0.744	0.744	0.000	0.002	0.001
ln_ADMEXPR	0.724	0.744	0.744	0.000	0.002	0.001

[a] R^2 from pooled cross-sectional regression of Equations (5) (cost equation) and (6) (test score equation).
[b] Omitted variable is the dependent variable or instrumental variable in the regression.
All variables are defined in Exhibit 1.

variables for *CSINS* replacing the original single variable. The results of these regressions are reported in Table 7.

For the sake of brevity, only the regression statistics for the three partitioned variables for *CSINS* are reported in Table 7 (along with the original variable in the pooled sample). As the results in Table 7 show, the results for *CSINS_1* and *CSINS_2* are consistent with the prior findings of a DB (double-benefit) variable. That is, for the first two groups, increased spending on instruction is associated with both lower overall spending and improved test scores. However, for the third partition (where the instructional cost share is above 62.2 percent), the situation reverses, with higher spending on instruction associated with both higher costs and lower test scores. This unexpected finding suggests that rule-of-thumb prescriptions that call for increased diversion of resources to hire more teachers, without due attention to their proper deployment, the development of appropriate curriculum, etc., can have adverse consequences. This kinked cost-behavior, whereby the benefits

TABLE 7
Tests of Linearity Assumption Regarding Cost Share of Instruction (CSINS) Pooled Cross-Sectional Results 2000–2002

	Cost Equation			Test Score Equation			
Variable CSINS[b]	**Coefficient**	**t-value**	**Significant Level**	**Coefficient**	**t-value**	**Significant Level**	**Relative Cost Effectiveness[a]**
CSINS Unpartitioned	−0.678	−13.58	.0001	0.065	4.25	.0001	DB
CSINS Partitioned[c]							
CSINS < 59.2%	−0.502	−6.01	.0001	0.086	3.51	.0005	DB
59.2% < CSINS < 62.2%	−0.524	−7.04	.0001	0.084	3.85	.0001	DB
CSINS >62.2%	0.566	8.51	.0001	−0.077	−3.94	.0001	DN

[a] Relative Cost Effectiveness Key:
DB = doubly beneficial (i.e., beneficial effects on both costs and test score outcomes).
DN = doubly negative (i.e., negative effect on both costs and test score outcomes).

[b] The distribution of CSINS in the combined sample is as follows: Minimum = 47.04 percent, Median = 60.8 percent, and Maximum = 73.2 percent.

[c] One-third of the observations of CSINS in the pooled sample (2000–2002) are between 59.2 percent and 62.2 percent.

of additional spending on instruction are lost, may explain why there are conflicting findings reported in the educational literature on the importance of direct spending on instruction (Hanushek 1986), and deserves further study.

Additivity Effects

No attempt was made in our main results to determine whether our findings vary cross-sectionally between high- and low-income school districts. In this section, we divide the sample into two sub-samples: (1) low-income districts (districts where the variable *HIGHINC* is below the median of 0.92) and (2) high-income districts (districts where *HIGHINC* is above the median). Equations (5) and (6) were re-estimated for these two sub-samples. Only the results for the outcome equation are reported in Table 8 (the cost function results are omitted), but the last two columns apply our methodology to show how the two sub-samples differ with respect to the cost-effectiveness of the management tools.

We focus the discussion on the last two columns of Table 8, where the cost-effectiveness labels on the independent variables differ between high-income and low-income districts. There are five variables with inconsistent cost-effectiveness ratings. Of these five, two are uncontrollable factors (*LEP* and *HIGHINC*). *LEP* has a trade-off rating (TO) for low-income districts, but a benefits only (BO) effect for high-income districts. This relation is reversed for *HIGHINC*. Greater incomes within the low-income district sub-sample are associated with higher test scores but not lower costs. But in the high-income district sub-sample, a trade-off is observed with respect to *HIGHINC* (higher test scores but also greater costs).

Among the controllable factors, three variables have differing signs on cost-effectiveness (*STUCOMP*, *FACEXPR*, and *CSADM*). Increasing the proportion of students to computers in the low-income sub-sample reduces costs with no effect on test score outcomes (BC). In contrast, in the high-income sub-sample, *STUCOMP* has a trade-off effect (TO), with lower test scores offset by lower spending. The level of faculty experience has a decreasing effect on student achievement in the low-income sub-sample with no offsetting cost reduction (NO), but no effect within the high-income sub-sample (NS). Our finding for the low-income sub-sample contradicts the results of some prior studies that found a positive or no relation between teacher experience and achievement scores (see for example, Elliott 1998; Hanushek 1986).[26] This finding may in part be explained by Hanushek (1986, 1162), who posits that student achievement affects teacher selection. In this case, it may be that experienced teachers with better teaching skills select schools with higher-achieving students residing in high-income districts. Lastly, greater costs shares devoted to administration has a beneficial effect on total costs with no offsetting in reduction in test scores (BC) in the low-income sub-sample, but a double benefit (both cost-reducing and test performance-increasing) in the high-income districts. This result is contrary to the results of Dee (2005), who finds a negative relation between non-instructional per-pupil expenditures and high-school graduation rates, and Brewer (1996), who finds no significant relation between number of school district administrators and achievement scores. Our finding on how changing cost shares for different functions can have such marked effects on costs and outcomes represents an avenue of future research.

We performed additional robustness tests to check for possible interaction effects. For example, it could be argued that teacher experience may be beneficial to test score outcomes in the poorer school districts (e.g., the Abbott districts), even if not significant overall. Likewise, lower student-faculty ratios could be more beneficial to test scores in districts

[26] It should be noted that the prior works cited do not partition their school samples on measures of median family income, making comparisons difficult with regards to teacher experience.

TABLE 8
Results of Sensitivity Tests for High and Low Income School for Test Score Equation Pooled Cross-Sectional Results 2000–2002

	Low Income Schools (*HIGHINC* < 92%)[b]		High Income Schools *HIGHINC* < 92%		Net Effect on Costs and Outcomes[a]	
Variable	Coefficient	t-value Significant Level	Coefficient	Significant Level	Low Income	High Income
Intercept	−0.865	−7.94***	−0.845	−5.76***	—	—
ELEM	0.071	10.32***	0.037	6.62***	—	—
MIDD	0.003	1.20	0.012	3.45***	—	—
HIGH	−0.008	−2.37*	0.006	1.01	—	—
ln_EXPPP	0.048	6.69***	0.051	4.75***	—	—
ln_WDFG	0.055	11.24***	0.021	6.23***	TO	TO
ln_GEOCEI	−0.090	−4.06***	−0.100	−3.57***	DN	DN
ln_STMOB	−0.005	−3.97***	−0.013	−4.52***	NO	NO
ln_SPED	−0.004	−1.07	0.005	1.06	CI	CI
ln_LEP	0.003	3.83***	0.005	4.24***	TO	BO
ln_HIGHINC	0.276	5.21***	0.074	10.88***	BO	TO
ln_STUFAC	0.009	1.59&	0.022	2.49*	DB	DB
ln_STUCOMP	−0.001	−0.62	−0.007	−2.55*	BC	TO
ln_ATTD	0.171	1.59&	0.395	3.61***	BO	BO
ln_FACSAL	0.008	1.10	−0.012	−1.01	CI	CI
ln_CSINS	0.040	2.34*	0.077	3.30***	TO	TO
ln_FACEXPR	−0.015	−4.35***	0.004	0.71	NO	NS
ln_ADMSAL	0.020	2.87**	0.043	4.00***	TO	TO
ln_CSADM	−0.006	−0.18	0.015	2.04*	BC	DB
ln_ADMEXPR	−0.007	−2.49*	−0.011	−2.73**	NO	NO

*, **, ***, **** Significant at 0.05 percent, 0.01 percent, 0.001 percent, and 0.10 percent levels, respectively.
[a] Relative Cost Effectiveness Key:
BC = beneficial to costs, but no effect on test score outcomes;
BO = beneficial to test score outcomes, but no effect on costs;
DB = doubly beneficial (i.e., beneficial effects on both costs and test score outcomes);
DN = doubly negative (i.e., negative effect on both costs and test score outcomes);
CI = cost increasing, with no off-setting beneficial effect on outcomes;
NO = negative effect on test score outcomes, but no effect on costs;
TO = trade-off effect (i.e., either a beneficial effect on costs offset by negative effect on test score outcomes, or vice versa); and
NS = no statistically significant effect on costs or outcomes.
[b] Median value of *HIGHINC* in pooled sample (2000–2002) is 92 percent.
All variables are defined in Exhibit 1.

marked by high student mobility. Neither of these two possible interaction effects were significant when estimated. We also surveyed the educational economics literature for examples of interaction terms estimated by other researchers, but failed to find evidence of interactions among the variables used in these studies examined. Given the large number of possible interaction terms that could be conceived of, without a theory to guide why particular interaction terms should be expected, further tests for potential interactions among variables did not seem practical.

CONCLUSIONS

In conclusion, the technique presented in this paper for identifying cost-effective management tools has the potential to yield valuable insight into the trade-offs made by the management of public-sector institutions with their multiple constituencies and objectives. The methodology proposed does not require any skills beyond some statistical knowledge, and the ability to identify and quantify some of the key management tools that may be applicable to a given context. Our empirical illustration of the methodology using public school data has demonstrated that the technique is feasible in an actual decision-making setting. By applying this methodology in a public education setting, school officials can identify factors controllable by school district superintendents to institute reforms to improve their operations and achieve greater cost efficiency and effectiveness.

Among the limitations of this study, perhaps the most important is that the empirical illustration draws on the public school data for only one state. As noted earlier, the results shown here may not be generalizable to other states due to institutional constraints and taxpayer preferences. Another potential limitation is that the illustration uses only one type of public-sector institution, public schools. Clearly, the application of the methodology to public school data in other states and to other types of public-sector institutions would help to establish the wider applicability of the suggested technique. Finally, the method proposed in the paper assumes linear additivity of the independent variables. To the extent that there are significant interaction terms, this assumption may be invalid, limiting the general applicability of the proposed method.

REFERENCES

Baber, W. R., P. L. Daniel, and A. A. Roberts. 2002. Compensation to managers of charitable organizations: An empirical study of the role of accounting performance measures of program activities. *The Accounting Review* 77 (3): 679–693.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley & Sons.

Bound, J., D. Jaeger, and R. Baker. 1995. Problems with instrumental variables estimation when correlation between the instruments and the endogenous explanatory variables is weak. *Journal of the American Statistical Association* 90 (430): 443–450.

Brewer, D. 1996. Does more school district administration lower educational productivity? Some evidence on the "administrative blob" in New York public schools. *Economics of Education Review* 15 (2): 111–124.

Broom, C. A. 1995. Performance-based government models: Building a track record. *Public Budgeting & Finance* 15 (Winter/4): 3–17.

Brown, R. E., M. J. Myring, and C. G. Gard. 1999. Activity-based costing in government: Possibilities and pitfalls. *Public Budgeting & Financ*e 19 (Summer/2): 3–21.

Callan, S. J., and R. E. Santerre. 1990. The production characteristics of local public education: A multiple product and input analysis. *Southern Economic Journal* 57 (2): 468–480.

Childress, S., R. Elmore, and A. Grossman. 2006. How to manage urban school districts. *Harvard Business Review* (November): 55–68.

Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, and R. York. 1966. *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Dee, T. 2005. Expense preference and student achievement in school districts. *Eastern Economic Journal* 31 (1): 23–44.

Dopuch, N., and M. Gupta. 1997. Estimation of benchmark performance standards: An application to public school performance. *Journal of Accounting and Economics* 23 (2): 141–161.

Elliott, M. 1998. School finance and opportunities to learn: Does money well spent enhance students' achievement? *Sociology of Education* 71 (3): 223–245.

Gianakis, G. A. 2002. The promise of public sector performance measurement: Anodyne or placebo? *Public Administration Quarterly* 26 (Spring): 35–64.

Governmental Accounting Standards Board (GASB). 1989. *Research Report, Service Efforts and Accomplishments Reporting: Elementary and Secondary Education*. Norwalk, CT: GASB.

———. 1994. *Service Efforts and Accomplishments Reporting (April)*. Concepts Statement No. 2. Norwalk, CT: GASB.

———. 1998. *Accounting and Financial Reporting for Nonexchange Transactions (December)*. Statement No. 33. Norwalk, CT: GASB.

———. 1999. *Basic Financial Statements—and Management's Discussion and Analysis—for State and Local Governments (June)*. Statement No. 34. Norwalk, CT: GASB.

Greene, W. 1980. Maximum likelihood estimation of econometric frontier production functions. *Journal of Econometrics* 13 (1): 27–56.

———. 1990. A gamma-distributed stochastic frontier model. *Journal of Econometrics* 46 (1–2): 141–163.

Hahn, J., and J. Hausman. 2002. A new specification test for the validity of instrumental variables. *Econometrica* 70 (1) (January): 163–189.

———, and ———. 2003. Weak instruments: Diagnosis and cures in empirical econometrics. *The American Economic Review* 93 (2): 118–125.

Hansmann, H. 1996. *The Ownership of Enterprise*. Cambridge, MA: Belknap Press.

Hanushek, E. 1979. Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources* 14 (3): 351–388.

———. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24 (3): 1141–1177.

———. 1989. The impact of differential expenditures on school performance. *Educational Researcher* 18 (4): 45–62.

———, J. F. Kain, and S. G. Rivkin. 2004. Disruption and Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics* 88 (9–10): 1721–1746.

Hausman, J. 1978. Specification tests in econometrics. *Econometrica* 46 (6): 1251–1271.

Hedges, L., R. Laine, and R. Greenwald. 1994. Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher* 23 (3): 5–14.

Jaggia, S., and A. Kelly-Hawke. 1999. An analysis of the factors that influence student performance: A fresh approach to an old debate. *Contemporary Economic Policy* 17 (2): 189–198.

Johnston, J. 1972. *Econometric Methods*. 2nd edition. New York, NY: McGraw-Hill.

Kinney, W. 2001. Accounting scholarship: What is uniquely ours? *The Accounting Review* 76 (2): 275–284.

Kravchuk, R. S., and R. W. Schack. 1996. Designing effective performance: Measurement systems under the Government Performance and Results Act of 1993. *Public Administration Review* 56 (4): 348–358.

Larcker, D. F., and T. O. Rusticus. 2005. On the use of instrumental variables in accounting research. SSRN Working paper. Available at: http://ssrn.com/abstract=694824.

Librera, W. 2003. Designation of Abbott districts: Criteria and process. New Jersey Department of Education (April 11). Available at: http://www.nj.gov/njded/abbotts/regs/criteria.htm.

McNab, R. M., and F. Melese. 2003. Implementing the GPRA: Examining the prospects for performance budgeting in the federal government. *Public Budgeting & Financ*e (Summer/2): 73–95.

Mensah, Y. M., and S.-H. Li. 1993. Measuring productive efficiency in a not-for-profit setting: An extension. *The Accounting Review* 68 (1): 66–92.

Mosso, D. 1999. Accounting for the business of government—New goals, old myths. *Public Budgeting & Finance* 19 (Winter/4): 65–74.

Mullins, D. R., and C. K. Zorn. 1999. Is activity-based costing up to the challenge when it comes to privatization of local government services? *Public Budgeting & Finance* 19 (Summer/2): 37–58.

Mundlak, Y. 1996. Production function estimation: Reviving the primal. *Econometrica* 64 (March): 431–438.

New Jersey Department of Education (NJDOE). 2004. *Vital Education Statistics 2004–2005*. Available at: http://www.state.nj.us/njded/data/vitaled/0405/.

Rose-Ackerman, S. 1996. Altruism, nonprofits, and economic theory. *Journal of Economic Literature* 34 (June): 701–728.

Sander, W. 1993. Expenditures and student achievement in Illinois: New evidence. *Journal of Public Economics* 52 (3): 403–416.

Shephard, R. 1970. *Cost and Production Functions*. Princeton, NJ: Princeton University Press.

Staiger, D., and J. Stock. 1997. Instrumental variable regression with weak instruments. *Econometrica* 65 (3): 557–586.

Stock, J., J. Wright, and M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics* 20 (4): 518–529.

Tiebout, C. M. 1956. A pure theory of local expenditures. *The Journal of Political Economy* 64 (5): 416–424.

Toch, T. 2003. *High Schools on a Human Scale: How Small Schools Can Transform American Education*. Boston, MA: Beacon Press.

U.S. Department of Education, Office of Educational Research and Improvement. 1990. *Financial Accounting for Local and State School Systems*. Washington, D.C.: National Center for Education Statistics.

U.S. Department of Education. 2002. No Child Left Behind Act of 2001. Public Law 107–110. Available at: http://www.ed.gov/policy/elsec/leg/esea02/index.html.

Varian, H. 1992. *Microeconomic Analysis.* 3rd edition. New York, NY: Norton.

Willoughby, K. G. and J. E. Melkers. 1998. The state of the states: Performance-based budgeting in 47 out of 50. *Public Administration Review* 58 (Spring/1): 66–73.

———, and ———. 2000. Implementing PBB: Conflicting views of success. *Public Budgeting & Finance* 20 (Spring/1): 105–120.