Pergamon

Economics of Education Review 19 (2000) 27–45

Economics of Education Review

www.elsevier.com/locate/econedurev

The ineffectiveness of school inputs: a product of misspecification?

James Dewey [a], Thomas A. Husted [b], Lawrence W. Kenny [c,*]

[a] *Bureau of Economic and Business Research, University of Florida, Gainesville, FL 32611, USA*
[b] *Department of Economics, American University, Washington, DC 20016, USA*
[c] *Department of Economics, University of Florida, Gainesville, FL 32611, USA*

Received 12 February 1997; accepted 23 October 1998

Abstract

Two-thirds of education production function studies relating learning to school and parental inputs also include parental income. This confounds demand and production functions, since in demand functions income determines the school inputs used in the production function. In a comprehensive review of the literature, we show that with this misspecification significantly positive school input coefficients are 39% less common. Then, with Project TALENT student-level data from 1960 and pooled state data for 1987–1992, we examine the impact of including income with no other change in specification. This causes most school inputs to become less significant. Hausman tests suggest that in OLS regressions there is a correlation between the school input measures and the error term, perhaps due to the omission of a good measure of parental time with the student. This appears to bias the school input coefficients toward zero, but can be corrected with IV methods. [JEL I21, H52, H42] © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Educational economics; Productivity

1. Introduction

Many academics seem to have accepted the conclusion of Hanushek (1986, p. 1162) that "*there appears to be no strong or systematic relationship between school expenditures and student performance*" (italics in original). Hanushek's rather startling statement concludes his review of 65 regressions, found in over a dozen articles, estimating how school expenditures affect educational outcomes. Although this conclusion is now widely accepted by economists, voters continue to ratify sizable educational budgets and spend more on education when they have more income. Real spending per pupil has tripled since 1960. Given this preference for increased spending, perhaps it is premature to conclude that school inputs are ineffective. In this paper, we reexamine the educational production function literature, focusing on the role of specification in determining the significance of education inputs.

Estimates from an educational production function are most commonly used to assess the efficacy of primary and secondary school inputs on educational output. This production function relates a measure of educational output (e.g., standardized achievement test scores, graduation rates, or post school economic achievements) to a set of inputs. The parental, student, and school inputs appropriate for these studies must reflect the fact that before and during the primary and secondary years knowledge is produced both at home and at school. Following standard production function theory, each input is expected to boost learning. Children should acquire more knowledge when their parents spend more time with them developing their skills, and more educated parents are expected to be more effective in passing on knowledge. Similarly, students should learn more when they spend more time in class learning from teachers,

* Corresponding author. Tel.: +352-392-0117; Fax: +352-392-7860; E-mail: Kenny@notes.cba.ufl.edu

PII: S0272-7757(99)00015-1

at home being tutored by their parents, or studying by themselves. Students are expected to learn more from better teachers and in smaller classes, where each student gets more attention. Though it may be less obvious why there should be higher achievement in larger schools, Kenny (1982) shows that the cost of any level of achievement is minimized by expanding the area served by a school until the increase in transportation costs per student is equated to the savings in in-school costs that accompany a larger school. This reduction in in-school costs implies that schools *should* operate in a region of increasing returns to scale and accordingly that students *should* learn more in larger schools, holding other inputs constant.[1]

The lack of support for these predictions in the education production function literature may be caused by misspecification. In two-thirds of the studies that we analyze below, family income or socioeconomic status has been included as an independent variable. Perhaps this stems from a belief that family factors "tend to be highly correlated with socioeconomic status of the family" (Hanushek, 1971, p. 281) and therefore either income or socioeconomic status can be used as a proxy for any inputs which families might provide outside of the school environment, but which are difficult to identify and measure. In the usual household production model, however, the parental time and income budget constraints determine the time parents spend with each child and quality of the school system.[2] In a Tiebout (1956) setting, parents select school inputs through the choice of the community; parents demanding better school locate in school districts that provide better schools. Accordingly, it makes no sense to include family income in an educational *production function* along with the parental and educational inputs that are jointly determined by family income. Doing so confounds production and demand functions.

We show in the next section that the inclusion of the *demand* variable 'family income' creates serious statistical problems that should make it more difficult to conclude that school inputs are significant. A few researchers have recognized this specification problem. Strauss and Sawyer (1986, p. 42), noting an "extensive public finance literature [in which] income [would be] a causal factor in the expenditure equation," conclude that educational production functions with income and educational expenditure mix a "demand-for-public-goods relation…and a production relation." They examine the impact of deleting income from the estimated production function. Unfortunately, their regressions do not include measures of parental or student time inputs. Gyimah-Brempong and Gyapong (1991) also recognize the endogeneity of these school inputs. They agnostically estimate the effect of dropping four socioeconomic measures—family income, parental educational attainment, poverty rate, and crime rate—from their regressions.[3] Their results suggest that multicollinearity between income and school spending leads to a negative coefficient on spending when income is included and a positive coefficient when income is excluded. Further, they find that dropping parental education results in a statistically significant loss of explanatory power in all cases, while all of the others can be dropped with no loss if parental education is retained.[4]

Although these studies, particularly the Gyimah-Brempong and Gyapong paper, suggest that including income may contribute to concluding that school inputs are ineffective for student achievement, there is no systematic evidence on whether this form of model misspecification is a problem. We provide evidence on this issue, using a variety of statistical techniques. In Section 3, we summarize, using meta-analysis, the empirical findings from a larger set of papers than Hanushek (1986) examined. Each regression studied is classified as either correctly specified (i.e., a 'good' study) or misspecified (i.e., a 'bad' study), and results for the two types of studies are compared. In Section 4, we estimate two educational production functions, using the Project TALENT micro data set and pooled state-level data, to determine directly the sensitivity of our parameter estimates to various specification errors involving the inclusion of income. The results of the meta-analysis and our regressions support our assertion that misspecification has played an important role in concluding that school inputs do not matter. We also use an instrumental variables procedure to deal with potential bias of the school input coefficients and find evidence of an endogeneity problem and that school input coefficients in even the better specified education productions appear to be biased downward.

2. Statistical effects of inclusion of demand variables along with inputs

To illustrate the specification issues, consider the following simple household production model. The parent's

[1] Although a number of papers provide support for this prediction, other research finds that costs fall and then rise as schools become larger. This literature is summarized in Cohn and Geske (1990, pp. 205–207).

[2] See, for example, Kenny (1982) and similar models in DeTray (1973), Willis (1973), and Becker (1991).

[3] It makes little sense to drop educational attainment, which measures effective parental time.

[4] In a related vein, Hanushek (1996, p. 21) claims that school input-earnings studies are marred by the omission of any parental input measures, which leads to overstatement of the impact of school inputs.

utility depends on how much the child learns (L) and other consumption (C).

$$U = U(L, C) \quad (1)$$

Each is produced in a household production function.

The education production function is

$$L = f(S, t_L; E) \quad (2)$$

where S represents the quality of the school system, t_L the amount of time the parent spends teaching the child, and E the educational attainment of the parent. An increase in the parent's education is expected to raise the productivity of his or her time spent teaching the child.[5] This formulation reflects the reasonable belief that virtually all of the parent's direct impact on the child's learning occurs through the quality and quantity of time spent with the child.[6]

Similarly, the production function for other consumption is

$$C = g(X, t_C; E) \quad (3)$$

where X is the amount of goods purchased and t_C is the time devoted to other consumption.

The parent faces the time and goods budget constraints:

$$T = t_L + t_C + t_W \quad (4)$$

where t_W is the time spent working and T is available time.

$$P_S S + X = wt_W \quad (5)$$

where P_S is the price of school quality in terms of X and w is the hourly wage rate. Substituting Eq. (4) into Eq. (5) produces the full income constraint:

$$P_S S + X = w(T - t_L - t_C) \quad (6)$$

$$(P_S S + wt_L) + (X + wt_C) = wT$$

This shows how potential income is allocated to the time and goods inputs into learning and other consumption.

Maximizing utility Eq. (1) subject to the production functions Eq. (2) and Eq. (3) and the full income constraint Eq. (6) yields the demand functions for the two outputs and their inputs:

$$L = D1(T, w, P_S, E)$$

$$C = D2(T, w, P_S, E)$$

$$S = D3(T, w, P_S, E)$$

$$X = D4(T, w, P_S, E)$$

$$t_L = D5(T, w, P_S, E)$$

$$t_C = D6(T, w, P_S, E)$$

Note that all outputs and inputs are determined by the same set of variables.

Now we can consider the issues involved in estimating the education production function Eq. (2). Suppose we first linearize it (ignoring error terms for the time being).

$$L = b_0 + b_1 S + b_2 t_L + b_3 E \quad (7)$$

We are aware of no data set that has ideal measures of all the variables entering the education production function. A few data sets examining the skills of young children have qualitative information on the time the mother spends teaching the child, but do not have good information on the child's school. The data sets with good information on schools have no direct measures of how much time the parents spend teaching their children (t_L). If a parental labor force measure (such as t_W) is available, it can be used as an instrument for t_L in Eq. (7), since it is clear from the time budget constraint (Eq. (4)) that these tend to be negatively correlated. The Project TALENT regressions that we report in Section 4 use this strategy.

Suppose more typically that there is no parental labor force variable. There are several ways in which to deal with this problem. Most commonly, the production function is estimated using a measure of the parent's income, such as w, presumably as a proxy for t_L. Then

$$L = b_0 + b_1 S + b_3 E + b_4 w \quad (8)$$

The problem with this solution is that w is unlikely to be a good proxy for the parent's time teaching the child. Conflicting income and substitution effects accompany an increase in the wage rate, making the net effect on t_L quite uncertain.

Moreover, the introduction of w into the production function introduces multicollinearity between w ('income') and the school quality S that it determines via the demand function $D3$. Since income is an important determinant of school quality, this problem will be severe. As multicollinearity may already be present in these models, the estimators may lack precision even without the inclusion of income. Adding income then might doom one to finding school inputs to be irrelevant, even

[5] There is substantial evidence that those who are more educated are more productive or more efficient. The labor market acts as if this is true in the work place, since wage rates rises with education. Michael (1973) found that the effect of education on the demand for goods, holding income constant, was consistent with a model of household production in which education raised household productivity neutrally. Those who are more educated are more adept at dealing with disequilibria (Schultz, 1975), are better informed about the voting record of their Senators (Husted, Kenny & Morton, 1995) and more likely to not reelect Senators who have been too liberal or conservative for their constituency (Schmidt, Kenny & Morton, 1996).

[6] Purchased household goods account for a very small share of the total cost of learning. Murnane, Maynard and Ohls (1981, p. 369) found no evidence that these goods affect achievement.

when they are in fact highly productive. Inspection of the expression below for the variance of the *j*th coefficient under standard ordinary least squares assumptions will make the severity of this problem apparent. Here, σ^2 is the variance of the error term and R_j^2 is the R^2 from an auxiliary regression of X_j on all of the other explanatory variables.

$$\mathrm{Var}[\beta_j] = \frac{\sigma^2}{(1 - R_j^2)\Sigma_i(X_{ij} - \bar{X}_j)^2}$$

Obviously, the inclusion of income will increase R_j^2, probably by a large amount. An analogous result may be obtained for the *estimated* variance as well.[7] Inclusion of income thus will increase the dispersion of estimates obtained from repeated regressions and increase the estimated standard errors in each individual regression. This means that more negative coefficients and fewer significant coefficients will tend to be found when income is included. The subsequent analysis provides evidence on these assertions.

Alternatively, the production function can be estimated without a measure of how many hours the parent spends with the child:

$$L = b_0 + b_1S + b_3E \tag{9}$$

The major econometric concern over this solution is the bias in the school quality coefficient due to omitting t_L. Flyer and Rosen (1997) found a positive correlation between female labor force participation rates and the teacher/pupil ratio, suggesting a negative correlation between school quality and parental time spent with the child and a negative bias on the school quality coefficient.

A third strategy is to use an instrumental variable procedure to estimate the demand function for school quality (*D*3) and then use the predicted values for school quality (S') to estimate a production function that does not include a measure of parental hours.

$$L = b_0 + b_1S' + b_3E \tag{10}$$

This procedure, followed in Kenny (1980) and in Section 4 below, should produce unbiased school quality coefficients. Its success, of course, hinges on using good instruments for school quality.

3. Meta re-analysis of education production function literature

3.1. Study selection and classification

To examine the production of knowledge before the child leaves the nest and ensure comparability, we confine our analysis to studies that examine cognitive skills, as measured by a score on a standardized test. Thus, studies examining earnings and high school completion are excluded from our analysis, but included in some other literature reviews. Hypotheses about various school inputs are confirmed in some studies, but not others. To allow testing for overall significance, we require that the result appear in a table in which the *t*-ratio was reported or from which could be calculated.[8] Thus, anecdotal statements that a variable was insignificant in unreported regressions were ignored; as Hedges, Laine and Greenwald (1994b, p. 9) note, in these situations “even the specification of the model was often unclear.” We omitted these ‘results’ because the level of significance could not be determined and furthermore because it is unclear how including the variable would have affected the significance of those variables that were included in the regression. Many studies report more than one regression. If a paper reports several regressions using the same dependent variable and sample, only the ‘best’ regression is utilized. On the other hand, regressions based on different dependent variables (for example, math and reading scores) and/or samples (e.g., based on grades or geographical area) provide complementary statistical evidence. Accordingly, each of these contributes potentially one observation to our analysis. The tables that follow are based on 127 regressions taken from 46 papers. Note that the analysis of Hanushek (1986) is based on 147 regressions from 33 publications. Our study includes the 28 papers he surveyed which included regressions meeting our criteria as well as 18 other papers.

These papers examined the impact of a variety of educational inputs. We collected results for teacher experience, teacher education, other teacher characteristics (e.g., teacher test score, ranking of teacher’s college), teacher salary, teachers per pupil, expenditure per pupil, and school size. If a regression included several variables from one category, such as science class size and non-science class size, the result for each variable was collected.

Classification of the regressions as correctly or incorrectly specified was based upon inclusion of appropriate variables to measure the impacts of each type of input. Many studies did not include any measure of the time students devoted to learning. However, we did not use this as a sorting criterion because such data have not commonly been available. All of the studies we examine consider the impact of school resources. The actual

[7] See Greene (1993, pp. 248, 267) for a thorough presentation of these issues.

[8] Similarly, if a variable and its square appeared in the regression and the estimated effect changed sign within the range of the sample, the results were excluded. In this case, it was impossible to ascertain significance from the reported results.

screening then centered upon two criteria related to family inputs. The first reason for labeling a regression 'bad' was outright failure to include some direct measure of parental contribution, such as parent's education and/or parent's hours worked. The second reason was inclusion of income or socioeconomic status, which as noted above are demand variables instead of input measures.

Under these two weak criteria (failure to include a measure of parental inputs and inclusion of income or socioeconomic status), 92 of the 127 studies, nearly three-quarters, were labeled as bad. Income or socioeconomic status was included in 85 studies; seven of these also did not have a variable specifically measuring parental inputs. Seven other studies were labeled as bad due to the omission of any parental input measure.[9] Misspecification, particularly the inclusion of income as a regressor, thus appears to be a serious problem in this literature.

3.2. Analysis of results

The 414 coefficients from the 127 regressions are classified according to sign and significance level in Table 1. The reported significance levels are based on a one-tail test; given the presumption that inputs have positive marginal products, such a test is appropriate. In Tables 2 and 3, our results are further summarized by sign and significance level, respectively, and compared with Hanushek's summaries (Hanushek, 1986).

If school inputs were truly ineffective, only half of the coefficients would be positive. The fact that over two-thirds (68.8%) of our coefficients, collected for the re-analysis, have the expected positive sign suggests that school inputs have a positive impact. Over 81% of the coefficients for expenditure per pupil and for other teacher characteristics (e.g., teacher test scores) are positive, but less than half the coefficients for school size are so. Hanushek, on the other hand, found only 56.1% of the coefficients to be positive in the studies he examined. The coefficients on teacher education, teachers per pupil, and expenditure per pupil were much less often positive in his summary than in the group of papers we examined.

There is generally little difference in coefficient signs between the studies we classify as good and those we classify as bad. In aggregate, 75.2% of the good studies and only 66.7% of the bad studies have the correct sign. Three quarters of this difference is attributable to the school size coefficients, which are nearly twice as likely to be positive in the good studies. Other teacher characteristics also are more commonly positive in correctly specified studies.

Thirty-seven percent of the 414 coefficients are significantly positive at a 5% (one-tail) significance level. To facilitate comparison with Hanushek (1986), Table 3 reports the frequency with which the coefficients are significantly positive at the 2.5% level with a one-tail test (or, equivalently, at the 5% level using a two-tail test). Almost a third of our coefficients are significantly positive, far more than the 2.5% that would be due to chance. There is considerable variation in the frequency of significance. Nearly half the coefficients on teacher experience, other teacher characteristics, and per pupil spending are significant, but only a sixth of the coefficients for teacher education, teachers per pupil, and school size are significant.

Once again, Hanushek's summary was less favorable to educational inputs than ours. Only 20.3% of his coefficients were significant. For each of the five input categories, his coefficients were less often significant than were the coefficients in our re-analysis. The biggest difference was for spending; his were half as likely as our coefficients to be significant. How much of this is due to our slightly different selection and recording criteria and how much is due to the inclusion of 18 new papers is not certain. However, the fact that only five of the papers yielding data in Hanushek's survey were dropped and 18 new ones were included suggests that the latter factor is dominant.

Specification is important to conclusions about statistical significance. Good studies were 39% more often significant than were bad studies. The correctly specified studies were much more likely to find that school size and various indicators of teacher quality were significant. Specification did not matter for conclusions about the effect of smaller class size; few good or bad studies found that to matter. Surprisingly, spending was significant more often in bad studies than in good studies. The meager effect of specification on sign, together with the sizable impact on significance suggests that misspecification produces insignificant results largely through higher estimated errors.

Describing the frequency with which a variable is highly significant, called vote counting, provides some information about the variable's statistical success, but does not result in a conclusion about its overall significance in the group of studies. Furthermore, "even when an effect is present in every study, vote counting typically has very low power to detect effects (it is prone to type II errors)" (Greenwald, Hedges & Laine, 1994, p. 10). We followed the meta-analysis literature and used the inverse chi-square test, known also as the Fisher test or the Pearson P_λ test, to assess overall significance.[10]

[9] As Hanushek (1996, p. 21) notes, the importance of school inputs should be overstated in studies with no parental measure. Nevertheless mixing these seven studies that favor education unduly with the 85 that we argue are biased against school inputs does not affect our empirical findings.

[10] See Maddala (1977, pp. 47–48) and Hedges and Olkin (1985).

Table 1
Meta re-analysis: coefficient summaries, showing sign and significance levels

	Positive			Negative		
1-tail significance	< 2.5%	2.5–5%	Insignificant	Insignificant	< 5%	Total
A: All 127 regressions						
Teacher education	11	7	25	12	9	64
Teacher experience	27	2	15	15	5	64
Teacher salary	7	3	3	6	3	22
Other teacher characteristics	33	5	28	8	2	76
Teachers per pupil	11	5	30	10	6	62
Expenditure per pupil	21	1	13	4	4	43
School size	15	4	19	20	25	83
Total	125	27	133	75	54	414
B: 35 good regressions						
Teacher education	3	0	5	3	1	12
Teacher experience	12	0	3	5	3	23
Teacher salary	0	0	0	0	0	0
Other teacher characteristics	10	1	7	1	0	19
Teachers per pupil	3	2	7	3	2	17
Expenditures per pupil	4	1	6	1	1	13
School size	8	0	7	4	2	21
Total	40	4	35	17	9	105
C: 92 bad regressions						
Teacher education	8	7	20	9	8	52
Teacher experience	15	2	12	10	2	41
Teacher salary	7	3	3	6	3	22
Other teacher characteristics	23	4	21	7	2	57
Teachers per pupil	8	3	23	7	4	45
Expenditures per pupil	17	0	7	3	3	30
School size	7	4	12	16	23	62
Total	85	23	98	58	45	309

Let p_i be the probability that the coefficient is, say, less than or equal to zero (the null hypothesis). Assuming that each coefficient represents an independent test of the null hypothesis, then it can be shown that $\Sigma - 2\log_e p_i$, $i = 1,2,...k$, has a χ^2 distribution with $2k$ degrees of freedom. This statistic will be used to test the null hypothesis that the parameter is less than or equal to zero, based on evidence from the set of k regressions. The non-positive null could be rejected by many moderately significant or a few highly significant positive coefficients. This test has been used by Kenny (1980), Greenwald, Hedges and Laine (1994), Hedges, Laine and Greenwald (1994a), Hedges and Greenwald (1996), and others to evaluate various education production function results.

One drawback of our analysis is that one-tail tests such as the Fisher test can yield seemingly inconsistent results. That is, it is possible to reject the null hypothesis that the coefficient is non-positive in all studies and also reject the null hypothesis that the coefficient is non-negative in all studies. Much of this is due to a small number of highly significant coefficients with the unexpected sign. Economic theory, of course, implies that inputs always should have a positive coefficient. To eliminate the influence of a few aberrant studies, we follow common practice in this literature and drop the 5% of the results that were most significantly positive and the 5% of the results that were most significantly negative; 370 of the original 414 coefficients remain. Inconsistency, not a major problem, was virtually eliminated by culling the results in this fashion.

As noted above, the inverse chi-square test is based on the assumption that these coefficients are independent. Since results taken from several studies that use the same sample and dependent variable are not independent, we use the mean probability across the common-data studies which remain after deleting the extreme positive and negative 5% to summarize the significance of the school input in that data set with that achievement test. Thus, only one probability is entered for each sample and dependent variable in the calculation of the χ^2 statistic.[11] This solution to common-data studies reduces the number of ‘results’ from 370 to 333.

The inverse chi-square tests based on these results are

Table 2
Meta analysis: comparison of the frequency of positive signs in Hanushek's study and our study and comparison of the frequency of positive signs in studies classified as good and in studies classified as bad

	Percent with positive coefficients			
		Dewey, Husted, and Kenny meta re-analysis		
	Hanushek[a]	All	Good	Bad
Teacher education	46.4	67.2	66.7	67.3
Teacher experience	69.1	68.8	65.2	70.7
Teacher salary	66.7	59.1		59.1
Other teacher characteristics		86.8	94.7	84.2
Teachers per pupil	37.4	74.2	70.6	75.6
Expenditure per pupil	70.4	81.3	84.6	80.0
School size		45.8	71.4	37.1
Total	56.1	68.8	75.2	66.7

[a]Hanushek (1986). Coefficients listed as having unknown sign are excluded.

Table 3
Meta analysis: comparison of frequency of significantly positive signs in Hanushek's study and our study and comparison of frequency of significantly positive signs in studies classified as good and in studies classified as bad

	Percent with significantly positive coefficients[a]			
		Dewey, Husted, and Kenny meta re-analysis		
	Hanushek[b]	All	Good	Bad
Teacher education	8.7	17.2	25.0	15.4
Teacher experience	35.1	42.2	52.2	36.6
Teacher salary	25.0	31.2		31.2
Other teacher characteristics		43.4	52.6	40.4
Teachers per pupil	9.9	17.7	17.6	17.8
Expenditure per pupil	24.1	48.9	30.8	56.7
School size		18.1	38.1	11.3
Total	20.3	30.2	38.1	27.5

[a]Significant at the 2.5% level using a one-tail test.
[b]Hanushek (1986). Coefficients listed as having unknown sign are excluded.

reported in Table 4. Panel A gives the evidence from both good and bad studies on the coefficient being positive, where the null hypothesis is that the coefficient is non-positive. The first and fourth columns indicate the number of 'results,' and the second and fifth columns report the probability of accepting the null hypothesis. It is rejected for all school input categories for the good studies and for all categories but school size for the bad studies. Based on this test alone, there is strong evidence from both correctly and incorrectly specified tests that school inputs boost learning.

Evidence that the coefficient is negative is reported in Panel B. The good studies accept the null hypothesis that the coefficient is non-negative for the coefficients as a group and for all input categories but school size. The incorrectly specified studies reject the null for the group as a whole and for school size and accept it for the other categories.

Combining the results from both panels, there is strong evidence from both good and bad studies that each school input except school size has a positive impact on learning. The results from the good studies do not provide consistent evidence on whether the coefficients on school size are positive. The misspecified studies are even less favorable to the hypothesis that students learn more in larger schools; they suggest that the true coefficient is negative.

Our analysis is based on a larger literature than was summarized by Hanushek. With this broader set of results, we find less inconsistency with the inverse chi-square test than did Hedges, Laine and Greenwald (1994a) in their analysis of the papers summarized by Hanushek. In our full set of 414 regression results, there were (unreported) inconsistent results on teacher education and school size. Hedges, Laine and Greenwald did not summarize school size results and report inconsistencies on teacher education, teacher salary, and the pupil/teacher ratio for the full sample. With the independent, middle 90% set of results reported in Table 4, we find only the school size results in the good studies to be inconsistent. Hedges, Laine, and Greenwald using similar procedures find inconsistency for teacher salary and teacher education.

4. Estimating production functions: sensitivity to specification

In this section we estimate educational production functions, based on two distinct data sets, to determine

[11] We already have mitigated this problem by using only the 'best' regression for a dependent variable and sample from any paper. We also tried using the median rather that the mean to combine common-data studies. This did not change any of the conclusions from the analysis; in fact the results were nearly identical. We also tried combining all results from the same sample using different dependent variables so that each data set would yield only a single observation. However, this left too few observations among the good studies to provide any meaningful insight.

Table 4
Meta re-analysis: inverse chi-square tests, based on middle 90% of results, testing significance of school inputs in group of good studies and in group of bad studies[a]

	Good studies		Bad studies	
	#	Joint significance	#	Joint significance
	(1)	(2)	(3)	(4)
A: Positive case (H_0: $\beta \leq 0$)				
Teacher education	10	0.003	41	0.000
Teacher experience	17	0.000	34	0.000
Teacher salary	0		19	0.000
Other teacher characteristics	7	0.000	48	0.000
Teachers per pupil	15	0.000	35	0.000
Expenditure per pupil	11	0.000	21	0.000
School size	20	0.000	55	0.392
Total	80	0.000	253	0.000
B: Negative case (H_0: $\beta \geq 0$)				
Teacher education	10	0.939	41	0.563
Teacher experience	17	0.476	34	0.999
Teacher salary	0		19	0.645
Other teacher characteristics	7	0.999	48	0.999
Teachers per pupil	15	0.839	35	0.976
Expenditure per pupil	11	0.997	21	0.999
School size	20	0.000	55	0.000
Total	80	0.178	253	0.000

[a]The 5% of the results that were most significantly positive and the 5% that were most significantly negative were deleted.

the effect on the signs and significance of school inputs due to model misspecification. Three types of misspecification are considered.

Although a number of micro (student) level data sets contain reasonable proxies for the time parents devote to their children's education, many education production function studies include income together with these time proxies and other socioeconomic variables. With the Project TALENT micro data set from 1960, we determine the impact of including income with no other change in specification. Such a clean comparison generally is not possible with meta-analysis. Since the Project TALENT data set was used in a number of studies reviewed above, revisiting this data set has special relevance for evaluating the education production literature.

With aggregate-level data, reasonable proxies for parents' time inputs are more difficult to obtain. Using state-level data for 1987–1992, we estimate educational production functions, based on several different school input specifications, to test for the effects of including income in the production function regressions. We also estimate the impact of using family income instead of parents' education, which may occur when parent's education is unavailable.

Concern over bias in the school input coefficients arises from their correlation with omitted parental time. With the state-level data, an instrumental variable procedure is employed to ascertain the extent of this bias. The instrumental variable procedure is not used for the Project TALENT data set, since it has seven measures of the quantity and quality of parental time inputs.

Some comparison of these data sets is warranted. The Project TALENT data come from an era when there was a stronger relationship between income and school inputs, which would create a greater misspecification problem than would be observed with recent micro data. On the other hand, each school is represented by only seven students on average in our Project TALENT sample, which produces some error in estimated community characteristics for the school and thus less correlation between school-level income and school inputs than would otherwise be the case. Husted and Kenny (1998) show that recent efforts to equalize education spending have made schools less efficient; this would make it more difficult to find school inputs significant with the modern pooled state level data set used in this study. Hanushek, Rivkin and Taylor (1996) claim that aggregate studies suffer from specification bias due to the omission of state regulations and other characteristics of the state's education structure. What they fail to realize is that this problem also plagues the many micro studies that use large multi-state data sets.

4.1. Project TALENT production function

4.1.1. Sample and variables: ordinary least squares

The first data set is drawn from the Project TALENT data set, a stratified random sample of all U.S. students in grades nine through 12 in 1960. We use a subsample of nearly 4300 12th grade males with complete information on relevant variables. The data come from a battery of tests the students took and from remarkably detailed questionnaires filled out by both the student and his high school principal. The dependent variables are (1) a verbal composite (C-003) that measures information about literature, vocabulary, spelling, capitalization, punctuation, English usage, and effective expression, and (2) a mathematics composite (C′004) that measures information about mathematics, arithmetical reasoning, and introductory and advanced high school math.

The quantity and quality of parents' time is measured by the number of jobs currently held by the father (FATHER'S JOBS), time spent in the labor force by the mother in the 1950s (MOTHER'S WORK), number of children in the family (CHILDREN), the student's birth order (BIRTH ORDER), the father's and mother's educational attainment (FATHER'S EDUC, MOTHER'S EDUC), and the student's race (BLACK).

Each student was asked to pick one of five categories that best described his family's income. Since the resulting categorical data very crudely measure family income, we have instead used the common logarithm of the summed median 1959 sex-specific earnings, for those working 50–52 weeks, in the parents' detailed occupations (LOG FAMILY INC).[12] To better approximate situations where more aggregate income data are used (perhaps because that is the only source of income data), we also use the school average *for our sample* of the log of family income (LOG AVE FAMILY INC); since the 4252 students in our sample come from 579 schools, these school-based means are based on only seven students on average. Since average income is more highly correlated with school inputs than individual income, using average income should have a more deleterious effect on showing that school inputs matter.

Several variables cover school inputs. The impact of large schools is captured by LOG SCHOOL SIZE, the common logarithm of the number of students in grades 9–12 in the school. Average teacher experience (TEACHERS' EXPER), the fraction of teachers with a masters degree (TEACHERS' EDUC), and average class size (CLASS SIZE) measure the quality and quantity of teacher inputs. The tracking of students according to ability groupings (TRACKING) is included to test whether this increases school effectiveness.[13]

As a student spends more time in class, there is an increase in both his time input and school inputs. Class time variables used in this study include days in the school year (SCHOOL YEAR), minutes in class in the school day (SCHOOL DAY), and self-reported data on the number of full semesters missed (MISSED SEMESTERS) and number of days absent from school the prior year (DAYS ABSENT). Hours spent studying each week (HOURS STUDYING) further measures the time the student devotes to learning. If changing schools disrupts learning, the number of times the student has changed schools since first grade (CHANGED SCHOOLS) should have a negative impact on achievement.

4.1.2. Regression results: ordinary least squares

Six regressions are reported in Table 5 for the two dependent variables. The first and fourth regressions do not include any income measure. Our measure of the student's family income (LOG FAMILY INC) is included in the second and fifth regressions, and our measure of mean family income in the school (LOG AVE FAMILY INC) is used in the third and sixth regressions. The fits are good for micro data, and most of our hypotheses are supported.[14]

Family income is used in education production functions as a proxy for parental time. Our regressions include seven variables that more directly measure the quality and quantity of parental time. If parental time in the labor force comes at the expense of time with their children, then FATHER'S JOBS and MOTHER'S WORK should have negative coefficients. Nine of the 12 coefficients are indeed negative, but only FATHER'S JOBS is significant, in the VERBAL regressions. The parental time allocated to each child is expected to fall as the number of children in the family rises, and CHILDREN has the predicted negative impact on achievement. Children of higher birth order typically are born when their parents are older and commanding higher wages; the negative coefficients on BIRTH ORDER suggest that children born later receive less help from their parents. The highly significant positive coefficients on the parental education variables support the hypothesis that more educated parents are more effective in teaching their children. The coefficients on BLACK are negative and are marginally significant in the MATH regressions;

[12] If the mother did not work, the earnings of those with her educational attainment were used.

[13] Note that there were serious problems with Project TALENT's measure of spending per pupil, which is not used in this study.

[14] The first and fourth regressions are very similar to a regression which was included in the analysis of the previous section.

Table 5
Project TALENT projection function OLS regressions (absolute *t*-statistics in parentheses)

Independent variables	Mean (S.D.)	Verbal			Math		
		(1)	(2)	(3)	(4)	(5)	(6)
Intercept		74.62	− 22.80	− 145.7	− 1.378	− 238.0	− 434.7
		(6.87)	(1.26)	(5.09)	(0.06)	(5.97)	(6.90)
Father's jobs	1.191	− 1.053	− 1.001	− 1.001	− 0.104	0.025	− 0.00074
	(0.494)	(2.01)	(1.92)	(1.92)	(0.09)	(0.02)	(0.00)
Mother's work	1.880	− 0.037	0.024	− 0.0060	− 0.141	0.0087	− 0.080
	(2.769)	(0.39)	(0.26)	(0.06)	(0.68)	(0.04)	(0.39)
Children	3.379	− 0.753	− 0.684	− 0.656	− 1.262	− 1.095	− 1.072
	(1.915)	(4.36)	(3.98)	(3.82)	(3.33)	(2.90)	(2.84)
Father's educ	11.414	0.995	0.770	0.856	2.166	1.618	1.892
	(3.159)	(9.92)	(7.30)	(8.48)	(9.82)	(6.99)	(8.51)
Mother's educ	11.573	0.798	0.425	0.608	1.962	1.057	1.589
	(2.478)	(6.34)	(3.10)	(4.79)	(7.09)	(3.51)	(5.68)
Black	0.016	− 4.169	− 4.307	− 4.836	− 16.01	− 16.34	− 17.32
	(0.063)	(1.01)	(1.05)	(1.18)	(1.76)	(1.81)	(1.92)
Birth order	1.858	− 0.573	− 0.645	− 0.630	− 0.931	− 1.105	− 1.042
	(1.336)	(2.35)	(2.65)	(2.60)	(1.74)	(2.07)	(1.95)
Log school size	2.922	2.217	1.989	1.030	6.257	5.704	3.922
	(0.426)	(2.23)	(2.01)	(1.03)	(2.87)	(2.63)	(1.79)
Teacher's exper	13.124	0.098	0.097	0.104	0.037	0.033	0.048
	(4.129)	(1.43)	(1.42)	(1.53)	(0.24)	(0.22)	(0.32)
Teacher's educ	0.599	1.019	0.905	0.111	6.936	6.658	5.149
	(0.276)	(0.77)	(0.69)	(0.08)	(2.39)	(2.31)	(1.78)
Class size	27.690	− 0.156	− 0.161	− 0.114	− 0.244	− 0.257	− 0.161
	(3.915)	(2.00)	(2.08)	(1.47)	(1.42)	(1.51)	(0.94)
Tracking	0.462	1.085	1.122	1.196	4.195	4.286	4.412
	(0.262)	(0.98)	(1.02)	(1.09)	(1.73)	(1.78)	(1.83)
School year	181.31	0.117	0.111	0.098	0.160	0.144	0.123
	(4.881)	(2.00)	(1.90)	(1.69)	(1.25)	(1.13)	(0.96)
School day	343.42	0.0091	0.0087	0.0079	0.025	0.024	0.023
	(49.20)	(1.65)	(1.58)	(1.44)	(2.07)	(2.00)	(1.89)
Missed semesters	0.128	− 3.261	− 3.201	− 3.151	− 6.078	− 5.932	− 5.861
	(0.542)	(6.77)	(6.68)	(6.59)	(5.74)	(5.64)	(5.57)
Days absent	6.085	− 0.173	− 0.175	− 0.180	− 0.602	− 0.608	− 0.616
	(5.381)	(3.55)	(3.62)	(3.72)	(5.63)	(5.71)	(5.79)
Hours studying	9.120	0.617	0.605	0.595	1.751	1.723	1.708
	(5.713)	(13.4)	(13.2)	(13.0)	(17.2)	(17.1)	(16.9)
Changing schools	1.508	− 0.301	− 0.306	− 0.403	− 1.083	− 1.095	− 1.284
	(1.681)	(1.91)	(1.95)	(2.57)	(3.13)	(3.18)	(3.72)
Log family inc	3.948		26.94			65.42	
	(0.084)		(6.68)			(7.39)	
Ave log family inc	3.948			58.43			114.9
	(0.042)			(8.31)			(7.42)
Adjusted R^2		0.1752	0.1836	0.1883	0.2081	0.2180	0.2181
Root Mean Squared Error		16.811	16.725	16.678	36.956	36.724	36.722
Number of observations		4252	4252	4252	4252	4252	4252

African-Americans may have lower test scores because their parents attended lower quality schools. In the regressions where they are included, the two income measures have a positive and significant impact on test scores.

Most of the hypotheses regarding school inputs are supported in the 'good' regressions (Eq. (1) and Eq. (4)). In contrast to many educational production function studies, we find that students learn more in larger schools, as Kenny (1982) predicted. Achievement is also found to be higher when students are taught in smaller classes. The evidence on which aspect of teacher quality is important is inconsistent. Students have higher verbal scores when taught by more experienced teachers and are more proficient in mathematics when taught by more educated teachers; in the other regression, these variables do not affect test scores. The positive and significant coefficient on TRACKING in the MATH regression is consistent with students learning more when they are placed in homogeneous 'ability' groupings; with tracking, fewer students are lost or bored. TRACKING is, however, insignificant in the VERBAL regression.

There is much speculation but little evidence on the efficacy of student time. The four variables capturing student time together with school inputs generally have the predicted effects. Achievement is higher when students spend more time each day in school; a longer school year produces higher verbal scores but has no impact on math performance. Students who have lost a semester or have been absent many days have lower test scores. Time spent studying at home or in study periods is very effective. HOURS STUDYING has a positive and highly significant impact on achievement. The significantly negative coefficients on CHANGED SCHOOLS suggest that there is some adjustment cost to changing schools.

Since there are seven measures of the quality and quantity of parental time in these regressions, there is no rationale for adding income to these regressions. We nevertheless examine the impact of doing so. The student's family income measure (LOG FAMILY INC) is used in the second and fifth regressions. Including family income generally makes it more difficult to conclude that school inputs have their predicted impact. LOG SCHOOL SIZE, TEACHERS' EXPER, TEACHERS' EDUC, SCHOOL YEAR, and SCHOOL DAY become less significant, but the *t*-statistic for CLASS SIZE becomes larger. In the VERBAL regression, use of LOG FAMILY INC leads to the rejection of the hypothesis that the school year is significant at the 2.5% level (with a one tail test) and makes the school day coefficient insignificant at the 5% level.

School inputs fare even worse when the more aggregate school-based measure of income (LOG AVE FAMILY INC) is added to the specification. This is expected, since school inputs are more collinear with average income in the school than with the family's income. In the VERBAL regression, the following important changes in one tail significance levels are observed: (1) LOG SCHOOL SIZE: 0.013–0.151; (2) CLASS SIZE: 0.023–0.079; (3) SCHOOL YEAR: 0.023–0.045; (4) SCHOOL DAY: 0.050–0.074. In the MATH regression, there were a number of instances where common significance thresholds (0.025, 0.05, 0.10) were passed: (1) LOG SCHOOL SIZE: 0.002–0.037; (2) TEACHERS' EDUC: 0.008–0.038; (3) CLASS SIZE: 0.077–0.173; (4) SCHOOL DAY: 0.019–0.029. In one exception to this pattern, TEACHERS' EXPER becomes more significant when community income is included.

The inclusion of income often makes it more difficult to conclude that school inputs are effective in part because it seems to bias the school input coefficients toward zero. Adding LOG FAMILY INC lowers the coefficients on school size, teacher experience, teacher education, school year, and school day by 7% on average and raises the coefficients on class size slightly. Adding instead the more aggregate income measure LOG AVE FAMILY INC, which is more correlated with school inputs, reduces the LOG SCHOOL SIZE, TEACHERS' EDUC, CLASS SIZE, SCHOOL YEAR, and SCHOOL DAY coefficients by an average of 33% but raise the TEACHERS' EXPER coefficients.

4.2. Pooled statewide production function

4.2.1. Sample and variables: ordinary least squares

The second production function we estimate is based on pooled statewide data, which have less measurement error than is often found in educational production function studies. Since there is much less mobility of students across states than across school districts or schools, the state educational characteristics more closely approximate the students' educational experiences than do current district or school educational measures. On the other hand, having better educational measures would be of little value if student and parental inputs were more poorly measured. In contrast to many other statewide studies, we have good information on the test-takers and their families.

Using ordinary least squares (OLS), we first estimate several specifications of the simple education production function traditionally found in this literature.[15] Our educational output measure in each is the average state total verbal and math SAT score. These test scores as well as several of our independent input variables are obtained from Educational Testing Service (ETS) reports on 37 states over 6 years. ETS did not publish reports on test-takers from the other 13 states due to the small numbers and percentages of students who took the SAT examinations in each year.[16] With the ETS reports, we are able to utilize the characteristics of the test-takers and their families instead of the characteristics of the general state

[15] We do not attempt here to estimate the effects of different state policies, which has been the subject of some recent research (see Hanushek, Rivkin & Taylor, 1996; Husted & Kenny, 1998).

[16] Since the thirteen omitted states—Idaho, Iowa, Kansas, Kentucky, Louisiana, Mississippi, Montana, Nebraska, North Carolina, North Dakota, South Dakota, Utah, and Wyoming—have a wide range of average SAT scores, omitting these states from our analysis is unlikely to be a problem.

population, found in many statewide studies.[17] Obviously, statewide characteristics do not always correspond with the characteristics of the state's test-takers. With test-taker data, we can better estimate the impact of family inputs and student characteristics on achievement. Our sample of 222 observations starts in 1987, the first year for which parental educational attainment was reported by ETS, and ends in 1992. To complete the basic specification, we combine these test-taker characteristics with state-level measures of public school characteristics. The summary statistics for the independent variables from these models are reported in Table 6.

In the first set of OLS regressions, we examine the basic forms of model misspecification. The first type that we consider arises when important family characteristics are unavailable and family income, usually some measure of state income, is substituted as a proxy measure for these missing characteristics. This form of misspecification is most prevalent when aggregate data like these are used to estimate the educational production functions. Because the ETS data include a measure of parent's education, it is possible to estimate the effects of this misspecification.

We measure parents' education by the fraction of test-takers' parents with an Associate's Degree or higher (PARENT'S EDUCATION). We use as our measure of income real state per capita personal income in 1982–1984 dollars (REAL PER CAPITA INCOME). In addition to these variables, we incorporate other family characteristics taken from the ETS data set. To allow for any historical inferiority of resources devoted to black education, we include the fraction of test-takers who are black (BLACK). Since educational systems often address the needs of the typical student, teaching may be less effective in states with more heterogeneous pupils. We test this hypothesis with a Gini coefficient of test-taker family income inequality (FAMILY INCOME GINI), which is calculated from the reported grouped data on income.

Table 6
Summary statistics for state production function variables

Independent variables	Mean	S.D.
Parents' education	4.0519	0.1600
Black	1.5982	0.9687
Family income gini	1.2142	0.0452
Public	4.3612	0.1035
English	4.3886	0.0500
Math	4.0994	0.1136
Large school	1.5259	2.2927
Pupils per teacher	2.8254	0.1195
Teachers' experience	3.6919	0.1461
Teachers' education	3.8569	0.2374
Teachers' rel wage	9.7780	0.1177
Spending per pupil	8.2086	0.2242
Real per capita income	9.5375	0.1483
Participation	3.4771	0.7600

To capture the students' exposure to the major SAT subjects, we include the fractions of test-takers who took four or more classes in English (ENGLISH) and math (MATH). We test the hypothesis that students in large schools learn more with the variable LARGE SCHOOL, which equals the percentage of test-takers who attend a school with 500 or more students in the senior class.

Public schools are much more regulated than private schools and should be less efficient than private schools if the regulations impose constraints on the production process. Public schools also may lose some of their best students to private schools. A number of studies, most notably Coleman, Hoffer and Kilgore (1982), have found that public school students have lower test scores than do private school students. Our variable PUBLIC equals the fraction of test-takers who attend public school.

A number of general *public* school characteristics are taken from the annual issues of the *Digest of Educational Statistics*. The quantity of teaching inputs devoted to each student is given by the student/teacher ratio in the state's schools (PUPILS PER TEACHER).[18] Teacher quality is captured alternatively by (1) the fraction of teachers with 10 or more years of teaching experience (TEACHERS' EXPERIENCE) and the fraction of teachers with at least a Master's degree (TEACHERS' EDUCATION), or (2) the ratio of teacher salaries to average earnings in the state (TEACHERS' REL WAGE). The latter measure follows Card and Krueger (1992), who normalize state teacher wages by the average earnings of all employees in the state covered by the social security system to control for geographic differences in cost of living, amenities, and in state-specific opportunity costs of teaching. Finally, SPENDING PER PUPIL in 1982–1984 dollars is used as a summary measure of school inputs.

Although often reported as performance measures of state school systems, average state SAT scores contain a potential bias. States in which a greater share of high school students go on to college, and thus take SAT or ACT tests, draw students from further down the distribution of skills, pulling down average test scores. Moreover, if a state's universities do not require the SAT for admission and instead require the ACT or some other

[17] Powell and Steelman (1984) and Graham and Husted (1993) also use data on the test-takers.

[18] Pupils per teacher measures class size with error, due to differences across districts in preparation periods, the use of teachers in non-teaching functions, etc.

test, few take the SAT test and many who do take it are applying to out-of-state universities. Since admission to out-of-state universities is often more selective, this effect also would cause the average SAT scores in the state to be higher in states with low participation rates.

Several recent empirical studies attempt to correct this potential bias by including a measure of state participation rates directly in the regressions. Dynarski (1987) includes the state participation rate in his empirical educational production function. Behrendt, Eisenach and Johnson (1986) and Graham and Husted (1993) develop an alternate model of selection that yields a specific functional form of the relationship between participation and state SAT scores.[19] In all cases, the state test taking participation rate has a significantly negative impact on statewide average examination scores. Our measure (PARTICIPATION) is the number of test-takers divided by the number of public and private high school graduates.[20]

4.2.2. Regression results: ordinary least squares

Table 7 reports the results of nine OLS regressions. Each of the nine regressions uses the state average SAT score as the dependent variable and includes a common set of explanatory variables. All variables are measured in logarithms. Each of three sets of regressions has a common form. Using the previous terminology, the specification of the first regression in each set [regressions Eq. (1), Eq. (4), and Eq. (7)] is considered to be 'good.' That is, PARENTS' EDUCATION is included and REAL PER CAPITA INCOME is excluded. The other two regressions in each set introduce misspecification. The second regression [Eq. (2), Eq. (5), and Eq. (8)] includes both PARENTS' EDUCATION and REAL PER CAPITA INCOME. Finally, in order to gauge the impact of the most egregious form of model misspecification on the various school input variables, REAL PER CAPITA INCOME replaces PARENTS' EDUCATION in the third regression in each set [Eq. (3), Eq. (6), and Eq. (9)]. The regressions also differ by the included public school input measures. In the first six regressions, school inputs are captured by teacher quantity (PUPILS PER TEACHER) and the three measures of teacher quality described above. Teacher skills are measured in the first three regressions by teacher education and experience (TEACHERS' EXPERIENCE, TEACHERS' EDUCATION) and in second three regressions by the relative salary of teachers (TEACHERS' REL WAGE). In regressions Eqs. (7)–(9), SPENDING PER PUPIL is the only school input variable. In all nine OLS regressions, we get a very good fit and most of our hypotheses are supported.[21]

In the six regressions where they are included, PARENTS' EDUCATION has the expected positive and statistically significant effect on student achievement. In all of the specifications, average test scores are lower in states with more black test-takers. In part, this result may be because black test-takers or their parents went through worse schools than whites. It also may reflect the generally lower quality of education in many southern states with large black populations. In addition, REAL PER CAPITA INCOME has significantly positive coefficients, consistent with its role as a proxy for parental inputs.

An increase in the Gini coefficient on the family incomes of test-takers (FAMILY INCOME GINI) is associated with greater income inequality. The negative and significant coefficients estimated on FAMILY INCOME GINI in all of the regressions support our hypothesis that the educational system is less efficient when the student body comes from a more diverse background.

The estimated coefficient on PUBLIC is negative and significant in eight of the nine models. We, like others, find that public school students have lower test scores than private school students. This may reflect the relative inefficiency of the public schools or a sorting of the better students into private school. The coefficients on the fractions of test-takers who have had four or more English courses (ENGLISH) or math courses (MATH), on the other hand, vary in sign and significance.

PARTICIPATION controls for differences among states in the fraction of seniors who take the SAT tests. Our results are consistent with other research, which finds that SAT scores fall as a larger fraction of the state's high school students take the test.

The focus of our research is on the impact of this specific type of model misspecification on the various school input variables. In our 'good' models [regressions Eq. (1), Eq. (4), and Eq. (7)], there is strong evidence that school inputs matter to student achievement. In regression Eq. (1), TEACHERS' EXPERIENCE and TEACHERS' EDUCATION have statistically significant positive impacts on SAT scores. The coefficient on TEACHERS' REL WAGE in regression Eq. (4) is positive and significant as well. In both these regressions, SAT scores are significantly lower in states with more pupils per teacher. SPENDING PER PUPIL, the summary school input measure, has a highly significant positive impact on learning in Eq. (7). The positive and significant coefficients estimated on LARGE SCHOOL support

[19] The specific details are outlined in Behrendt, Eisenach and Johnson (1986).

[20] Our correction, like others in this literature, adjusts for selection in the group of high school seniors in taking the SAT test but does not deal with those who dropped out earlier.

[21] Again, the first regression here was included in the analysis of Section 1.

Table 7
State pooled production function OLS regressions (absolute *t*-statistics in parentheses)

Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	6.987	5.951	5.829	7.075	6.181	5.905	6.257	5.974	5.965
	(44.4)	(31.6)	(22.7)	(49.2)	(32.5)	(22.7)	(43.6)	(33.6)	(24.6)
Parents' education	0.200	0.135	–	0.200	0.163	–	0.160	0.149	–
	(17.2)	(7.98)		(17.5)	(13.8)		(15.5)	(13.5)	
Black	− 0.010	− 0.012	− 0.014	− 0.009	− 0.010	− 0.013	− 0.010	− 0.010	− 0.012
	(8.32)	(10.7)	(9.48)	(7.63)	(9.21)	(8.83)	(9.20)	(9.59)	(8.88)
Family income gini	− 0.116	− 0.080	− 0.072	− 0.113	− 0.095	− 0.090	− 0.106	− 0.097	− 0.088
	(4.74)	(3.67)	(2.44)	(4.60)	(4.16)	(2.89)	(4.83)	(4.43)	(2.97)
Public	− 0.072	− 0.039	− 0.026	− 0.073	− 0.041	− 0.015	− 0.078	− 0.061	− 0.037
	(6.58)	(3.74)	(1.84)	(6.91)	(3.75)	(1.04)	(8.11)	(5.30)	(2.35)
English	− 0.066	− 0.033	− 0.055	− 0.037	− 0.002	0.002	0.018	0.023	0.023
	(2.36)	(1.31)	(1.63)	(1.41)	(0.07)	(0.05)	(0.75)	(0.97)	(0.70)
Math	− 0.011	0.009	0.046	− 0.026	− 0.007	0.032	0.008	0.012	0.030
	(0.87)	(0.84)	(3.07)	(2.06)	(0.60)	(2.00)	(0.75)	(1.12)	(2.09)
Large school	0.0022	0.0012	0.0007	0.0024	0.0014	0.0004	0.0021	0.0015	0.0011
	(4.29)	(2.55)	(1.02)	(4.77)	(2.84)	(0.56)	(4.56)	(3.15)	(1.64)
Pupils per teacher	− 0.056	− 0.033	0.0114	− 0.075	− 0.054	− 0.007	–	–	–
	(5.51)	(3.53)	(0.96)	(7.23)	(5.41)	(0.56)			
Teachers' experience	0.018	0.030	0.0488	–	–	–	–	–	–
	(2.31)	(4.31)	(5.20)						
Teachers' education	0.0084	0.0077	− 0.0030						
	(1.80)	(1.88)	(0.56)	–	–	–	–	–	–
Teachers' rel wage	–	–	–	0.040	0.017	− 0.008	–	–	–
				(2.62)	(1.21)	(0.43)			
Spending per pupil	–	–	–	–	–	–	0.053	0.041	0.030
							(10.2)	(6.05)	(3.29)
Real per capita income	–	0.079	0.141	–	0.067	0.134	–	0.033	0.100
		(8.09)	(12.0)		(6.51)	(10.7)		(2.63)	(6.25)
Participation	− 0.031	− 0.046	− 0.077	− 0.033	− 0.045	− 0.078	− 0.040	− 0.044	− 0.08
	(12.0)	(15.7)	(31.7)	(12.5)	(14.8)	(30.1)	(15.9)	(14.8)	(31.6)
Adjusted R^2	0.934	0.950	0.907	0.934	0.945	0.895	0.945	0.946	0.900
Number of observations	222	222	222	222	222	222	222	222	222

the prediction of Kenny (1982) that schools operate in a region of increasing returns to scale.

The regressions summarized in the remaining columns illustrate the statistical problems that result when family income serves its dubious role as a proxy for family inputs. Including income in these education production functions causes four school variables to become less significant and one variable (TEACHERS' REL WAGE) to lose significance. Replacing parents' education with average income has a devastating impact, with four school measures (LARGE SCHOOL, PUPILS PER TEACHER, TEACHERS' EDUCATION, and TEACHERS' REL WAGE) becoming insignificant, three of these even changing sign! The only public school input that does not do worse when income is included is teachers' experience. Contrary to expectations, the positive coefficients on TEACHERS' EXPERIENCE become slightly more significant when income is included in regressions Eq. (2) and Eq. (3). The evidence strongly suggests that the inclusion of income biases the school input coefficients toward zero. Only the TEACHERS' EXPERIENCE coefficients become larger, and the coefficients of the other five school measures fall by 32% on average when income is included. Replacing education with income biases the coefficients even more toward zero; those coefficients that retained their sign fell by 65% on average.

4.2.3. Sample and variables: instrumental variables

As described in the Introduction, one problem with using OLS to estimate these educational production functions if parental time or other inputs are not included in the regressions is that the school inputs may be correlated with the omitted inputs and thus with the regression residuals. An instrumental variable procedure offers one solution to the negative bias that is expected to result in the school input coefficients.

We use separate regressions, corresponding to the school input demand functions *D*3, to create instruments for the five public school input measures—PUPILS PER

TEACHER, TEACHERS' EXPERIENCE, TEACHERS' EDUCATION, TEACHERS' REL WAGE, and SPENDING PER PUPIL. We use a common set of demographic and political characteristics of the state population as determinants of the levels of these five educational inputs.[22] To capture income effects, we include the measure of real per capita state personal income (REAL PER CAPITA INCOME) that we used in the OLS regressions and the spread of income in the state. INCOME SPREAD equals the difference between the first and third quartile family incomes divided by the median family income. Since the political parties represent different segments of the income distribution, we include a measure of which party controls state government. DEMOCRATIC CONTROL equals one (negative one) if the Democrats (Republicans) control both the legislative and executive bodies and equals zero if control of state government is shared by the two political parties. Other socioeconomic variables that we include to reflect differences in demand for education are the percentages of the state's population who are black (PERCENT BLACK), aged 65 years and older (PERCENT ELDERLY), live in metropolitan areas (PERCENT METROPOLITAN), and have a college education (PERCENT COLLEGE). The percent of the state's population who own their homes (PERCENT OWNER) is included to capture the expected differences in tastes for education between owners and renters as well as to reflect the prevalence of the ability to itemize some costs of home ownership on federal (and some state) income taxes.

4.2.4. Regression results: instrumental variables

The coefficient estimates from the instrumental variable regressions are reported in Tables 8 and 9. The results from the regressions used to create the five public school input instrumental variables are reported in Table 8. These regressions explain between 18% and 61% of the variation in the five school input measures.

REAL PER CAPITA INCOME is the most consistently significant variable in these regressions. Its importance in explaining school inputs points to the potential econometric problems of incorrectly including income in a production function equation with these school inputs. The hypothesized demand relationships between income and the school inputs are mostly supported. As expected, holding other state demographic and political variables constant, higher state per capita income leads to lower pupil–teacher ratios, a higher proportion of more experienced and higher educated teachers, and larger real per pupil education spending. The only unexpected finding was the inverse relationship between state per capita income and the teacher's relative wage.

The second stage involves incorporating these instrumental variables into a standard educational production function explaining variation in SAT scores. The coefficient estimates from three specifications, differing by the included public school inputs, are reported in Table 9. The results for the family, subject preparation, and test participation variables are virtually identical to those for the comparable OLS regressions in Table 7 (regressions Eq. (1), Eq. (4), and Eq. (7)). Also, these regressions provide additional, although somewhat weaker, evidence that students learn more in larger schools.

Once again, we find that with correctly specified regressions, most of the school input measures have a significant impact on achievement. The use of instrumental variables has little effect on significance. TEACHERS' EXPERIENCE loses significance, but the *t*-statistic on TEACHERS' EDUCATION rises from a marginal 1.80 to 2.75.[23] The OLS coefficients on school inputs appear to be biased toward zero, as predicted. When the instrumental variables are used, the absolute coefficients on PUPILS PER TEACHER, TEACHERS' EDUCATION, TEACHERS' REL WAGE, and SPENDING PER PUPIL rise by 55%, 912%, 88%, and 34%, respectively. The instrumental variable procedure, however, produces larger standard errors on these coefficients. Finally, Hausman tests indicate that the school inputs are endogenous.[24]

5. Conclusion

The common academic perception that schools are ineffective is inaccurate. In part, this is due to literature reviews that are based on an old and somewhat incomplete literature. Our more comprehensive summary of the literature is more favorable than earlier reviews to the hypothesis that school inputs have positive marginal products. Over two-thirds of our coefficients on school inputs are positive and almost a third are significant at the 2.5% level. These far exceed the 50% and 2.5% figures, respectively, that would be associated with there being no relationship. Using a technique drawn from meta-analysis, we find that teacher education, teacher experience, teacher salary, other teacher characteristics, teachers per pupil, and expenditure per pupil each have

[22] Poterba (1997) included many of these same variables in his model of the determinants of government education spending.

[23] Interestingly, the same changes in significance are observed in unreported Project TALENT instrumental variable estimates.

[24] The predicted school inputs are a significant addition to OLS specifications 1, 4, and 7 in Table 7. The *F*-statistics are 9.71, 7.64, and 6.14, respectively. See Maddala (1992, pp. 510–513).

Table 8
State regressions creating the instrumental variables (absolute *t*-statistics in parentheses)

Independent variables	Mean (S.D.)	Pupils per teacher	Teachers' experience	Teachers' education	Teachers' rel wage	Spending per pupil
		(1)	(2)	(3)	(4)	(5)
Intercept		7.793	7.727	1.747	1.540	-5.440
		(8.29)	(6.10)	(0.76)	(2.33)	(3.64)
Family income	9.537	− 0.686	0.146	0.270	− 0.234	1.124
	(0.148)	(8.80)	(1.39)	(1.43)	(4.27)	(9.08)
Income spread	− 0.021	0.046	− 0.108	0.031	− 0.557	0.035
	(0.077)	(0.41)	(0.71)	(0.11)	(7.05)	(0.20)
Percent black	1.838	− 0.017	− 0.046	0.068	0.011	0.020
	(1.131)	(1.67)	(3.35)	(2.73)	(1.57)	(1.25)
Percent elderly	2.493	− 0.211	− 0.318	0.158	0.007	− 0.014
	(0.227)	(5.52)	(6.18)	(1.70)	(0.25)	(0.24)
Percent metropolitan	4.228	0.311	0.143	0.051	0.040	− 0.218
	(0.315)	(8.27)	(2.82)	(0.55)	(1.52)	(3.65)
Percent college	2.539	0.501	− 3.092	− 0.305	0.773	1.428
	(0.036)	(1.64)	(7.50)	(0.41)	(3.58)	(2.94)
Percent owner	4.175	− 0.109	0.646	− 0.098	− 0.288	0.056
	(0.087)	(1.01)	(4.42)	(0.37)	(3.77)	(0.33)
Democratic control	0.266	0.004	− 0.001	− 0.063	0.001	− 0.052
	(0.568)	(0.38)	(0.09)	(2.34)	(0.13)	(2.96)
Adjusted R^2		0.451	0.335	0.178	0.309	0.607
Number of observations		222	222	222	222	222

a significantly positive impact on test scores in the set of studies reviewed. Only the hypothesis that students learn more in large schools is unsupported.

Misspecification is a problem in this literature. Two-thirds of the studies we review include income along with the parental, student, and school inputs that are determined by income. The inclusion of this extraneous variable should render finding school inputs significant more difficult, and that is indeed what we find. Significantly positive coefficients are 39% less common in misspecified studies.

We conducted our own empirical analysis using the Project TALENT student-level data set from 1960 and pooled state data for 1987–1992. In regressions from both data sets that were not plagued by misspecification, there is evidence that each school input had an impact on achievement. This conclusion is not altered when instrumental variables replace the usual school input variables. However, when income was added to the OLS regression, most school input measures became less significant. Indeed, including income caused several variables to cross some critical significance threshold. In the Project TALENT data, this problem is more acute when school-based family income is used than when the student's family income is used; this is not surprising, since there is a stronger correlation between school inputs and the more aggregate income measure. In the state data, school input results suffer more from the use of income when income is used in lieu of parents' education than when income is simply added to the regression. The evidence also suggests that this misspecification biases the school input coefficients downward by 7–65%.

The Hausman tests suggest that in the OLS specification that predominates in this literature there is a correlation between the school input measures and the error term. This may be due to omitting some measure of the time parents spend teaching their children. One solution is to use an instrumental variable procedure such as that followed here. We find that the OLS school input coefficients are biased toward zero, which is consistent with the negative correlation between parental inputs and school inputs suggested by the research of Flyer and Rosen (1997).

We agree with Hanushek and others that it does matter how money is spent on education; simply spending more will not guarantee higher achievement in all cases. The evidence on productivity is stronger for some inputs than for others. Further, while we find very strong evidence of widespread and consistent positive effects, the number of negative coefficients present even in good studies suggests that perhaps some schools do not do a very good job of allocating their resources. Still, if one wishes to ask the policy question 'Do school inputs matter for the educational attainment of children?' the answer is a resounding yes.

Table 9
State production function instrumental variable regressions (absolute *t*-statistics in parentheses)

Independent variables	(1)	(2)	(3)
Intercept	6.481	6.870	5.935
	(22.76)	(44.92)	(32.05)
Parents' education	0.210	0.199	0.137
	(10.52)	(15.35)	(11.12)
Black	− 0.015	− 0.007	− 0.009
	(4.71)	(5.85)	(8.01)
Family income gini	− 0.116	− 0.077	− 0.071
	(3.01)	(2.67)	(2.89)
Public	− 0.034	− 0.057	− 0.048
	(1.71)	(4.89)	(4.45)
English	− 0.030	0.0078	0.031
	(0.62)	(0.27)	(1.13)
Math	− 0.023	− 0.027	0.022
	(1.10)	(1.93)	(1.90)
Large school	0.001	0.002	0.001
	(1.39)	(3.82)	(2.87)
Pupils per teacher[a]	− 0.088	− 0.115	–
	(3.36)	(6.17)	
Teachers' experience[a]	0.018	–	–
	(0.78)		
Teachers' education[a]	0.085	–	–
	(2.75)		
Teachers' rel wage[a]	–	0.075	–
		(2.21)	
Spending per pupil[a]	–	–	0.071
			(7.89)
Participation	− 0.032	− 0.032	− 0.045
	(7.17)	(10.47)	(14.52)
Adjusted R^2	0.848	0.925	0.936
Number of observations	222	222	222

[a]Instrumental variable.

Acknowledgements

We have benefitted from comments from Chunrong Ai, Julie Cullen, David Denslow, Genia Toma, and three anonymous referees.

Empirical studies used

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., Zellman, G., Sumner, G., & Thompson, V. M. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools.* Santa Monica, CA: Rand.

Bieker, R. E., & Anschel, K. R. (1973). Estimating educational production functions for rural high schools: Some findings. *American Journal of Agricultural Economics, 55*, 515–519.

Boardman, A. E., Davis, O. A., & Sanday, P. R. (1977). A simultaneous equations model of the educational process. *Journal of Public Economics, 7*, 23–49.

Borland, M. V., & Howsen, R. M. (1992). Student academic achievement and the degree of market concentration in education. *Economics of Education Review, 11*, 31–39.

Bowles, S. (1970). Toward an educational production function. In W. L. Hansen, *Education, Income, and Human Capital* (pp. 11–61). New York: National Bureau of Economic Research.

Brewer, D. J. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review, 12*, 281–292.

Brown, B. W., & Saks, D. H. (1975). The production and distribution of cognitive skills within schools. *Journal of Political Economy, 83*, 571–593.

Cohn, E. (1968). Economies of scale in Iowa high school districts. *Journal of Human Resources, 3*, 422–434.

Cohn, E., & Millman, S. D. (1975). *Input–output analysis in public education*. Cambridge, MA: Ballinger.

Cooper, B. S., Sarrel, R., Darvas, P., Meier, E., Alfano, F., Samuels, J., & Hienbuch, S. (1994). Making money matter in education: A micro-financial model for determining school-level allocations, efficiency, and productivity. *Journal of Education Finance, 20*, 66–87.

Delbertin, D. L. (1976). Estimating education production functions in rural and urban areas. *Southern Journal of Agricultural Economics, 8*, 31–35.

Deller, S. C., & Rudnick, E. (1993). Production efficiency in elementary education: The case of Maine public schools. *Economics of Education Review, 12*, 45–57.

Dugan, D. J. (1976). Scholastic achievement: Its determinants and effects in the education industry. In J. T. Froomkin, D. T. Jamison, & R. Radner, *Education as an industry* (pp. 53–83). Cambridge, MA: Ballinger.

Eberts, R. W., & Stone, J. A. (1984). *Unions and public schools*. Lexington, MA: D.C. Heath.

Eberts, R. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review, 7*, 291–299.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review, 13*, 1–17.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation, 28*, 465–498.

Gyimah-Brempong, K., & Gyapong, A. O. (1991). Characteristics of education production functions: An

application of canonical regression analysis. *Economics of Education Review, 10*, 7–17.
Hanushek, E. A. (1970). The production of education, teacher quality, and efficiency. In *Do teachers make a difference?* (pp. 79–99). Washington: U.S. Office of Education.
Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996). Aggregation and the estimated effects of school resources. Working paper. Revised (1996) *Review of Economics and Statistics, 78*, 611–627.
Heim, J., & Perl, L. J. (1974). *The educational production function: Implications for educational manpower policy*. Monograph Number 4. Ithaca, NY: Institute of Public Employment.
Henderson, V., Mieskowski, P., & Sauvageau, Y. (1976). *Peer group effects and education production functions*. Ottawa, Canada: Economic Council of Canada.
Husted, T. A., & Kenny, L. W. (1995). *Evidence from the States on the political and market determinants of efficiency in education*. Working Paper 94-95-10. Gainesville, FL: University of Florida.
Jencks, C. S. (1973). The quality of the data collected by the equality of educational opportunity survey. In F. Mosteller, & D. P. Moynihan, *On equality of educational opportunity* (pp. 473–512). New York: Random House.
Katzman, M. T. (1971). *The political economy of urban schools*. Cambridge, MA: Harvard.
Kenny, L. W. (1980). The effects of family and school on cognitive achievement. *Research in Population Economics, 2*, 99–113.
Kiesling, H. J. (1967). Measuring a local government service: A study of school districts in New York State. *Review of Economics and Statistics, 49*, 356–367.
Kiesling, H. J. (1976). A study of the relationship of instructional process and program organization to the success of compensatory education projects in California. In J. T. Froomkin, D. T. Jamison, & R. Radner, *Education as an industry* (pp. 249–282). Cambridge, MA: Ballinger.
Levin, H. M. (1970). A new model of school effectiveness. In: *Do teachers make a difference?* (pp. 55–78). Washington: U.S. Office of Education.
Levin, H. M. (1976). Concepts of economic efficiency and educational production. In J. T. Froomkin, D. T. Jamison, & R. Radner, *Education as an industry* (pp. 149–191). Cambridge, MA: Ballinger.
Link, C. R., & Ratledge, E. C. (1979). Student perceptions, I.Q., and achievement. *Journal of Human Resources, 14*, 98–111.
Margo, R. A. (1986). Educational achievement in segregated school systems: The effects of separate-but-equal. *American Economic Review, 76*, 794–801.
Maynard, R., & Crawford, D. (1976). School performance. In D. L. Bawden, & W. S. Harrar, *Rural income maintenance experiment final report*, vol. 6, part 2 (pp. 1–104). Madison, WI: Institute for Research on Poverty.
Michelson, S. (1970). The association of teacher resources with children's characteristics. In *Do teachers make a difference?* (pp. 120–168). Washington: U.S. Office of Education.
Michelson, S. (1972). For the plaintiffs—Equal school resource allocation. *Journal of Human Resources, 8*, 283–306.
Montmarquette, C., & Mahseredjian, S. (1989). Does school matter for educational achievement? A two-way nested-error components analysis. *Journal of Applied Econometrics, 4*, 183–193.
Murnane, R. J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.
Murnane, R. J., & Phillips, B. (1981). What do effective teachers of inner-city children have in common? *Social Science Researcher, 10*, 83–100.
Perl, L. J. (1973). Family background, secondary school expenditure, and student ability. *Journal of Human Resources, 8*, 156–180.
Ribich, T. I., & Murphy, J. L. (1975). The economic returns to increased educational spending. *Journal of Human Resources, 10*, 56–77.
Saunder, W. (1993). Expenditures and student achievement in Illinois. *Journal of Public Economics, 52*, 403–416.
Sebold, F. D., & Dato, W. (1981). School funding and student achievement: An empirical analysis. *Public Finance Quarterly, 9*, 91–105.
Stern, D. (1989). Educational cost factors and student achievement in Grades 3 and 6: Some new evidence. *Economics of Education Review, 8*, 149–158.
Strauss, R. P., & Sawyer, E. A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review, 5*, 41–48.
Summers, A., & Wolfe, B. (1977). Do schools make a difference? *American Economic Review, 67*, 639–652.
Winkler, D. R. (1975). Educational achievement and school peer group composition. *Journal of Human Resources, 10*, 189–204.

References

Becker, G. S. (1991). *A treatise on the family*. Cambridge, MA: Harvard.
Behrendt, A., Eisenach, J., & Johnson, W.R. (1986). Selectivity bias and the determinants of SAT scores. *Economics of Education Review, 5*, 363–371.
Card, D., & Krueger, A.B. (1992). Does school quality matter?

Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, *100*, 1–40.

Cohn, E., & Geske, T. G. (1990). *The economics of education*. Oxford: Pergamon. Reprinted (1998) Columbia, SC: Academic Advantage.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement: Public, Catholic, and private schools compared*. New York: Basic.

DeTray, D.N. (1973). Child quality and the demand for children. *Journal of Political Economy*, *81*, S70–95.

Dynarski, M. (1987). The Scholastic Aptitude Test: Participation and performance. *Economics of Education Review*, *6*, 263–273.

Flyer, F., & Rosen, S. (1997). The new economics of teachers and education. *Journal of Labor Economics*, *15*, S104–139.

Graham, A.E., & Husted, T.A. (1993). Understanding state variations in SAT scores. *Economics of Education Review*, *12*, 197–202.

Greene, W. H. (1993). *Econometric analysis*. New York: Macmillan.

Greenwald, R., Hedges, L.V., & Laine, R.D. (1994). When reinventing the wheel is not necessary: A case study in the use of meta-analysis in educational finance. *Journal of Education Finance*, *20*, 1–20.

Gyimah-Brempong, K., & Gyapong, A.O. (1991). Characteristics of education production functions: An application of canonical regression analysis. *Economics of Education Review*, *10*, 7–17.

Hanushek, E.A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review*, *61*, 280–288.

Hanushek, E.A. (1986). The economics of schooling. *Journal of Economic Literature*, *24*, 1141–1177.

Hanushek, E.A. (1996). Measuring investment in education. *Journal of Economic Perspectives*, *10*, 9–30.

Hanushek, E.A., Rivkin, S.G., & Taylor, L.L. (1996). Aggregation and the estimated effects of school resources (Revised). *Review of Economics and Statistics*, *78*, 611–627.

Hedges, L. V., & Greenwald, R. (1996). Have times changes? The relationship between school resources and student performance. In G. Burtless, *Does money matter? The effect of school resources on student achievement and adult success* (pp. 74–92). Washington: Brookings.

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

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Money does matter somewhere: A reply to Hanushek. *Educational Researcher*, *23*, 9–10.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic.

Husted, T. A., & Kenny, L. W. (1998). Evidence on the impact of state government on primary and secondary education and the equity-efficiency tradeoff. Working paper.

Husted, T.A., Kenny, L.W., & Morton, R.B. (1995). Constituent errors in assessing their senators. *Public Choice*, *83*, 251–271.

Kenny, L.W. (1980). The effects of family and school on cognitive achievement. *Research in Population Economics*, *2*, 99–113.

Kenny, L.W. (1982). A model of optimal plant size with an application to the demand for cognitive achievement and for school size. *Economic Inquiry*, *20*, 240–254.

Maddala, G. S. (1977). *Econometrics*. New York: McGraw-Hill.

Maddala, G. S. (1992). *Introduction to econometrics*. New York: Macmillan.

Michael, R. (1973). Education in non-market production. *Journal of Political Economy*, *81*, 306–327.

Murnane, R.J., Maynard, R.A., & Ohls, J.C. (1981). Home resources and children's achievement. *Review of Economics and Statistics*, *63*, 369–377.

Poterba, J.M. (1997). Demographic structure and the political economy of public education. *Journal of Policy Analysis and Management*, *16*, 48–66.

Powell, B., & Steelman, L.C. (1984). Variations in state SAT performance: Meaningful or misleading? *Harvard Educational Review*, *54*, 389–412.

Schmidt, A.B., Kenny, L.W., & Morton, R.B. (1996). Evidence on electoral accountability in the U.S. Senate: Are unfaithful agents really punished? *Economic Inquiry*, *34*, 545–567.

Schultz, T.W. (1975). The value of the ability to deal with disequilibria. *Journal of Economic Literature*, *13*, 827–846.

Strauss, R.P., & Sawyer, E.A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review*, *5*, 41–48.

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

Willis, R.J. (1973). A new approach to the economic theory of fertility behavior. *Journal of Political Economy*, *81*, S14–64.