

Evidence of Student Attendance as an Independent Variable in Education Production Functions

DOUGLAS J. LAMDIN
University of Maryland

ABSTRACT Most studies of student performance that use the production function or input-output approach do not consider student attendance as an independent variable. Data from Baltimore public elementary schools indicate that student attendance is positively and significantly related to standardized achievement test performance. Consistent with other studies, the importance of the socioeconomic status of the students and the lack of a positive influence of school input measures such as the teacher/pupil ratio and expenditure per pupil are apparent.

A large literature exists on the empirical relationship between inputs into the education process and the performance level of students. The usual focus of attention in such studies is whether policy variables—particularly the teacher/pupil ratio or expenditure per pupil—affect student performance. One potential policy variable that receives scant attention in this literature is student attendance. A recent exception is Caldas (1993), who found that attendance is positively and significantly related to student performance. (See Romer, 1993, for an article on the way the attendance of undergraduates influences their grades and the replies to Romer in the Summer 1994 issue of the same journal.) In their studies, Berger and Toma (1994); Brown and Saks (1975); Eberts, Schwartz, and Stone (1990); Fowler and Walberg (1991); Oates (1977); Sander (1993); Walberg and Walberg (1994); and Wendling and Cohen (1981) did not examine attendance. Given the reasonable presumption that higher attendance at the individual or aggregate level should be associated with a higher level of student performance, this lack of attention is curious.

Moreover, attendance is a policy variable worthy of attention insofar as resources can be and are devoted to improve it.

Some speculative explanations for the reason that attendance has been absent from analyses of student perfor-

mance can be offered. First, it is unlikely that attendance has been absent from analyses because the data are not available. Perhaps analysts believe, or have found but do not state, that attendance and socioeconomic status are highly correlated. If collinearity is present, the separate influence of each variable is empirically difficult to disentangle with regression analysis. Analysts may therefore omit the attendance variable from the analysis rather than omit the socioeconomic variable(s). Also, because most of these studies were cross-sectional and based on aggregated data, perhaps there was relatively little variation in attendance. It may therefore be difficult to find any statistically significant relationship between attendance and student performance, even if one truly exists. My goal in this study was to shed some further light on the attendance–performance relationship, using data from the public elementary schools of Baltimore, Maryland.

The Production Function Approach

Analysts use the production function approach and multiple regression analysis to examine the relationship between the output of the education process and the inputs into this process. Hanushek (1979, 1981, 1986, 1989, 1991) provided discussions of this approach. The term *input-output approach* may be more appropriate than *production function approach* because the nature of the education process does not comport with the production function of economic theory in many respects. For example, there is no universally accepted single output measure. The education process may be characterized by multiple outputs. Also, the implicit assumption of an optimizing producer allocating inputs may not be tenable in the case of education.

Test scores that measure cognitive performance are gen-

Address correspondence to Douglas J. Lamdin, Department of Economics, The University of Maryland–Baltimore, 5401 Wilkens Avenue, Administration Building, Room 629, Baltimore, MD 21228-5398.

erally used as the output measure. Using test scores as the output measure is believed to be most appropriate at the elementary school level, where cognitive skills are the primary focus (Hanushek, 1986).

Student Inputs and School Inputs

The input measures can be categorized as *student input variables* or *school input variables*. The student input variables include innate student ability, parental background, and the socioeconomic status of the student. The school input variables include the teacher/pupil ratio, expenditure per pupil, and the amount of education and/or experience of the teacher. School input variables are generally manipulable, whereas student input variables are not. Caldas (1993) and others alternatively labeled student input variables as *input* factors and school input variables as *process* factors. The level of aggregation of the data in studies has ranged from the individual student, to the school, to the school district. Depending upon the nature of the questions being asked, the preferred level of aggregation differs.

Previous Studies

There are two major findings in this literature. The first is that student input measures, usually proxied by socioeconomic measures of the student, consistently are shown to affect student performance. The second is that school input measures such as expenditure per pupil or teacher/pupil ratio (or the inverse, class size) are not consistently shown to have a positive affect. Some researchers have found that school inputs are positively related to student achievement; most, however, have not reported a significant positive relationship (Hanushek, 1981, 1986, 1989, 1991). Murnane (1991) provided a dissenting opinion regarding the lack of importance of school inputs, as did Hedges, Laine, and Greenwald (1994a, 1994b). Despite these protestations, however, the weight of the evidence supports the view that there is no relationship (Hanushek, 1994).

Generally absent in production function analyses is examination of attendance. It is an empirical question as to whether attendance matters. Moreover, as a policy variable, it is desirable to know whether there is some systematic relationship between attendance and student performance.

To make clear what is at issue, consider the general function

$$Q = f(\mathbf{X}_1, \mathbf{X}_2).$$

The variables Q , \mathbf{X}_1 , and \mathbf{X}_2 represent student performance, a student input measure (or vector of measures), and a school input measure (or vector of measures), respectively. Inclusion of attendance in the \mathbf{X}_1 vector is the primary issue here. One could argue that attendance, insofar as it is or can be affected by school policies, belongs in the \mathbf{X}_2 vector also. Although this is an intriguing semantic distinction, for the empirical purposes here, this distinction does not matter.

The Data: Variables and Descriptive Statistics

The data used in the empirical analysis were tabulated in a 1990 report by the Baltimore Citizens Planning and Housing Association (CPHA). Data for 107 public elementary schools in the city of Baltimore, Maryland, concerning characteristics of the school, the students, and measures of the performance of the students on standardized tests were reported. Ten schools do not have Grades K through 5. Because these 10 schools were not comparable to the others, they were eliminated from the analysis. All the reported results are based on the 97 remaining schools.

The political subdivision of the city of Baltimore is one school district. It is the 14th largest public school district in the United States (U.S. Department of Education, National Center for Education Statistics, 1990). The fact that the data are for one district has a desirable attribute. Empirical studies are often based on data from a large cross-section of school districts. Teacher quality and school curriculum could influence student performance. In such studies, however, these variables are often not measured reliably or at all. Use of school level data within a single district may mitigate this problem because, a priori, variations in teacher background and curriculum should vary considerably less within a district than among districts that cover a large geographic area.

The Dependent Variable

The dependent variable in the analysis was student performance on the California Achievement Test (CAT) in the spring of 1989. The CAT has a reading portion and a mathematics portion. Data were available for each school for Grades 1 through 5 within each school and aggregated for the entire school for the reading and mathematics portions. The CPHA study reported the percentage of students in each school above the national median score. Therefore, analyses could be conducted at both the school and grade level, using the reading scores, the mathematics scores, or an average of the two scores. The school was the level of analysis used. The percentage of students in each school above the national median was the dependent variable used in the regression analyses.

The Independent Variables

The independent variables were proxies for student and school inputs. Following the standard approach, the student input measure was a socioeconomic measure: the percentage of students who do not qualify to receive free lunch. The percentage of minority (non-White) students in each school was not available in the CPHA report but was obtained from the *Enrollment by Race/Ethnic Categories* publication (Baltimore, Maryland State Department of Education, 1990). This variable is often used as an additional student input measure in studies of this nature.

The following school input measures were used: the teacher/pupil ratio and the professional staff/pupil ratio. The latter ratio adds the number of principal(s), librarian(s), counselor(s), and other specialists to the number of classroom teachers used in the numerator of the teacher/pupil ratio. An alternative school input measure, namely, operating expenditure per pupil, was also available. It included the expenditure on instructional programs only (salaries, materials, and supplies). Building and maintenance expenses were not included in this measure.

Table 1 contains descriptive statistics for the variables. For the analysis of the questions at hand, the larger the degree of dispersion in the variables, the better. Although the data were for a single school district, the variables exhibited a remarkable range of values. The greatest expenditure per pupil was more than twice that of the lowest. The greatest average class size (the reciprocal of the teacher/pupil ratio) was twice that of the lowest. Note that the variable with the lowest relative degree of dispersion, as measured by the coefficient of variation, was attendance. Although attendance did vary across the sample, it did not vary as much as the other variables.

Further exploration of descriptive statistics in the form of the correlation matrix in Table 2 before estimating regression models, allows for examination of potential collineari-

ty problems. Not surprisingly, the teacher/pupil and professional staff/pupil ratios and expenditure per pupil were significantly correlated with one another. These correlations, however, were far from perfect. The sensitivity of the results to using these three different school input measures is examined in the regression analyses that follow. There was a strong negative correlation between each school input measure and student attendance. Also, there was a strong positive correlation between no free lunch and attendance. The latter was alluded to earlier as a potential source of a collinearity problem. The former correlation, however, is not so easy to explain. One possible explanation is that administrators may direct resources disproportionately toward schools with students with lower socioeconomic status (which is apparent in the correlation matrix), and those schools also have lower attendance rates.

More Correlations

Examined, but not reported in detail here, were (a) the correlations between the reading and mathematics scores within grade (i.e., first-grade reading and first-grade mathematics score, second-grade reading and second-grade mathematics score, and so on) and (b) the correlations between grades (i.e., first-grade reading and second-grade reading

Table 1.—Descriptive Statistics for Dependent and Independent Variables

Variable	<i>M</i>	<i>Mdn</i>	<i>SD</i>	Min	Max	CV
% students above median reading	45.4	44.0	13.9	16.0	80.0	0.31
% students above median mathematics	56.4	56.0	12.8	29.0	93.0	0.23
% reading and mathematics average	50.9	50.5	13.1	22.5	86.5	0.26
% no free lunch	40.5	33.0	23.0	7.0	92.0	0.57
Teacher/pupil ratio	0.035	0.033	0.006	0.024	0.051	0.16
Professional staff/pupil ratio	0.054	0.052	0.009	0.041	0.087	0.17
Expenditure/pupil ratio	2,468.5	2,411.2	427.5	1,690.5	3,850.0	0.17
% minority	79.1	99.0	33.7	0.6	100.0	0.43
% attendance	90.1	90.0	2.15	85.0	95.0	0.02

Note. The number of observations was 97.

Table 2.—Bivariate Correlation Matrix of the Dependent and Independent Variables

Variable	1	2	3	4	5	6	7	8
1. % students above median, reading								
2. % students above median, mathematics	.92							
3. % reading and mathematics average	.98	.98						
4. Teaching/pupil ratio	-.49	-.47	-.49					
5. Professional/pupil ratio	-.22	-.19	-.21	.37				
6. Expenditure per pupil	-.49	-.41	-.46	.49	.70			
7. % no free lunch	.74	.69	.73	-.50	-.16	-.53		
8. % minority	-.39	-.46	-.44	.25	-.06	.13	-.60	
9. Attendance	.61	.56	.60	-.38	-.32	-.50	.54	-.11

Note. A coefficient that exceeded 0.16, 0.19, and 0.23 in absolute value was significantly different from 0 at the 10, 0, and 1% level for a two-tailed test. The number of observations was 97.

score, first-grade reading and third-grade reading score, and so on, as well as the first-grade reading score and the second-grade mathematics score, the first-grade reading and the third-grade mathematics score, and so on). These correlations measure two relationships of interest: (a) the extent to which reading and mathematics scores are correlated within grades, and (b) the extent to which the students in different grades within a given school exhibit similar relative performance within the district. All of the possible correlations were not only positive, but also significant at the 1% level. For example, the first-grade mathematics score and the fifth-grade reading score had a correlation coefficient of .39; the fifth-grade mathematics score and the first-grade reading score had a .35 coefficient. The reading and mathematics score correlation for each grade revealed the expected strong positive relationship. In particular, the correlation coefficients for reading and mathematics scores for Grades 1 through 5 were .77, .82, .86, .74, and .65. With scores aggregated to the school level, the correlation between the reading and mathematics scores was .92. The strong correlations among the grade level test scores within the schools suggest that it would be appropriate to use either school-level or grade-level data.

Regression Analysis Results

In this section, I present the results of the regression analyses used to test for the relationship between attendance and performance. In doing so, we also examine the relationships between student performance and the other input measures.

I used three different dependent variables: the percentage of students in each school above the median reading score, the percentage in each school above the median mathematics score, and the average of these two measures. The dependent variables were regressed on the no-free-lunch and attendance variables and one of the three school input measures. I used separate regressions, each of which includ-

ed one of the three school input measures, to test for the sensitivity of the results to these different school input measures (see Tables 3, 4, and 5).

The coefficient on the attendance variable was positive and statistically significant at the 5% level for a one-tailed test in all nine of the specifications and significant at the 1% level for eight of the nine. This result strongly suggests that attendance does have a positive influence on student performance, other factors held constant.

One must be careful, however, in attributing this positive relationship entirely to attendance in a regression analysis such as this. To the extent that the attendance variable is a proxy for latent variables such as innate student motivation, or parental concern, or the ability of the teacher to stimulate or motivate students, the true influence of attendance per se is overstated. However, this same caveat applies as well to the socioeconomic measures, such as the no-free-lunch variable, discussed below. The nature of latent variables, and the inability of the analysts to measure such variables accurately, or at all, is inherent in this type of research. Although one can recognize this potential problem of interpretation, assessing its magnitude is not a simple task.

The coefficient on the socioeconomic measure was positive and significant at the 1% level across all specifications. In light of previous studies, this is not surprising. Of additional importance is that both the attendance and no-free-lunch variables were statistically significant, even though the two variables were significantly correlated with one another. A collinearity problem, which would be evidenced by "artificially insignificant" *t* ratios of these coefficients, was not apparent.¹

As a further check on collinearity, I examined the scaled condition index for each set of independent variables in the regression models. The values of this index were well below 30 (i.e., 18.2, 14.7, and 17.2); this result is further evidence of no serious collinearity problem. (See Belsley, 1991; Belsley, Kuh, & Welsch, 1980; Gujurati, 1988; Johnston, 1984; and Kennedy, 1992, for discussions of the condition index.)

Table 3.—Regression Results: Teacher/Pupil Ratio Used as School Input

Variable	Constant	Teacher/pupil ratio	No free lunch	Attendance	R ²	n
Reading	-120.2	-290.3 (-1.56) [-0.22]	0.32 (6.39) [0.28]	1.81 (3.63) [3.59]	.61	97
Mathematics	-73.0	-295.1 (-1.58) [-0.18]	0.28 (5.58) [0.20]	1.42 (2.84) [2.27]	.54	97
Combined	-96.6	-292.7 (-1.65) [-0.20]	0.30 (6.30) [0.24]	1.62 (3.40) [2.86]	.60	97

Scaled condition index = 18.2

Note. Dependent variable: % of students in the school above the national median score. The *t* ratios are reported in parentheses. The *t* ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36. The output elasticities, as defined in the text, are reported in brackets.

Table 4.—Regression Results: Professional/Pupil Ratio Used as School Input

Variable	Constant	Professional/pupil ratio	No free lunch	Attendance	R ²	n
Reading	-133.9	-49.0 (-0.46) [-0.06]	0.35 (7.46) [0.31]	1.86 (3.59) [3.69]	.61	97
Mathematics	-91.5	-24.3 (-0.23) [-0.02]	0.31 (6.60) [0.22]	1.52 (2.90) [2.42]	.53	97
Combined	-112.7	-36.6 (-0.36) [-0.04]	0.33 (7.38) [0.26]	1.69 (3.41) [2.99]	.59	97

Scaled condition index = 14.7

Note. Dependent variable: % of students in the school above the national median score. The *t* ratios are reported in parentheses. The *t* ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36. The output elasticities, as defined in the text, are reported in brackets.

Table 5.—Regression Results: Expenditure per Pupil Used as School Input

Variable	Constant	Expenditure per pupil	No free lunch	Attendance	R ²	n
Reading	-128.3	-0.0017 (-0.66) [-0.09]	0.34 (6.74) [0.30]	1.82 (3.49) [3.61]	.61	97
Mathematics	-99.9	0.0005 (0.18) [0.02]	0.31 (6.23) [0.22]	1.58 (3.00) [2.52]	.53	97
Combined	-114.1	-0.0006 (-0.25) [-0.03]	0.32 (6.81) [0.25]	1.70 (3.41) [3.01]	.59	97

Scaled condition index = 17.2

Note. Dependent variable: % of students in the school above the national median score. The *t* ratios are reported in parentheses. The *t* ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36. The output elasticities, as defined in the text, are reported in brackets.

The coefficients on the school input measures were nearly all negative. More often than not, however, the coefficients were statistically indistinguishable from zero, although the coefficient on the teacher/pupil ratio was significantly negative at the 10% level in all three models. It is generally presumed that more teachers per student would enhance student performance, as would more professional staff, or more expenditure per pupil. The counterintuitive empirical finding here is not without precedent. Hanushek (1986) reported that 57 of the 112 studies he surveyed found a negative relationship between student performance and the teacher/pupil ratio, with 14 of these statistically significant.

A caveat may be warranted. The variation in school resources evident in Table 1 may not be random. Table 2 contains evidence that additional resources may be directed toward schools with more students of lower socioeconomic status. The three school input measures were negatively related to the percentage of students that do not qualify for free lunch. Insofar as the allocation of school resources is

determined, in part, by student performance, the coefficients on the school input variables must be interpreted cautiously because of problems induced by the potential mutual causation. The standard statistical procedure to correct for simultaneous-equations bias is two-stage least squares. In the circumstances here, however, the variables that may affect the way in which school resources are allocated are the same variables that are expected to affect student performance. Therefore, the two-stage least squares procedure that would correct for such a simultaneous-equations bias cannot be used to remedy this because of the identification problem. (For a discussion of this, see Gujurati, 1988; Johnston, 1984; and Kennedy, 1992).

The relative impact of each independent variable on the dependent variable cannot, of course, be assessed directly via the magnitude of the respective regression coefficients. A unit-free measure, namely, the output elasticity, is necessary. This represents the percentage change in the dependent variable (*Q*) for a 1% change in the independent variable (*X*) with each variable evaluated at the sample mean.

Specifically, this is calculated as $(\partial Q/\partial X) (Q/X)$ with Q and X as defined in the generalized expression for the production function. The regression coefficient is $(\partial Q/\partial X)$. These output elasticity values are reported in the square brackets in Tables 3 through 5.² Using the reading score equation from Table 3 as exemplary, the output elasticities for teacher/pupil ratio, no free lunch, and attendance were -0.22, 0.28, and 3.59. Clearly, for a given percentage increase, attendance would have the largest relative impact on test scores. For example, a 5% increase in attendance from 90.1% (the sample mean) to 94.6% (a bit below the maximum observation) would increase the reading score from 45.4% to 53.5% above the median.

Minority Students Variable Added

The percentage of minority students in the school was added to each of the models reported in Tables 3 through 5. In the interest of space, the detailed results are not reported. The coefficients on the other variables were essentially unaffected. The coefficients on minority students were a bit curious. The coefficients were negative but statistically insignificant (all t ratios less than 0.5 in absolute value) in the case of reading scores. In the case of mathematics scores, the coefficient was negative and always significant at the 10% level for a two-tailed test. Correspondingly, the combined score equations revealed an intermediate result. This result suggests that, other things being equal, schools with a larger minority enrollment will have lower mathematics scores, but no discernible difference in reading scores.

Results With the Multiplicative Function

An alternative specification of a regression model is the multiplicative (aka the power function or log-linear) model, rather than the linear model. In an n -input case, the linear model is

$$Q = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n.$$

The multiplicative model analog is

$$Q = a_0X_1^{a_1} X_2^{a_2} \dots X_n^{a_n}.$$

To allow for the estimation of the exponents (the regression coefficients) with ordinary least squares regression, one must first transform all of the variables by the natural logarithm function before the standard regression estimation. That is, one estimates the transformed model:

$$\ln Q = \ln a_0 + a_1 \ln X_1 + a_2 \ln X_2 + \dots + a_n \ln X_n.$$

A useful property of this specification is that the coefficient estimates are simultaneously the output elasticity estimates.³ That is, each regression coefficient represents the percentage change in the dependent variable for a 1% change in the independent variable, other variables held constant.

The multiplicative function also may be conceptually preferred to the linear model because of two other mathematical properties that appear plausible in this context. With the multiplicative function, the marginal impact of each input (i.e., $[\partial X/\partial Q]$) depends on the levels of the other inputs. Also, the marginal impact can increase, decrease, or be constant as that input increases. With the linear model, in contrast, the marginal impact of each input is independent of the level of the other inputs and the marginal impact is constrained to be constant regardless of the level of that input. For a further discussion see (e.g., Gujurati, 1988, pp. 144-146, or Salvatore, 1993, chapters 4 and 6).

For comparison, the results obtained with the multiplicative specifications are shown in Tables 6 through 8. Qualitatively these results are similar to those in the analogous Tables 3 through 5. One difference is that although the other output elasticities differed only trivially between the two functional specifications, the output elasticities with respect to attendance were invariably higher in the logarithmic specifications.⁴

Summary and Conclusions

The results presented here suggest that the average level of attendance at a school does have a positive influence on

Table 6.—Regression Results: Teacher/Pupil Ratio Used as School Input (All Variables Were Converted to the Natural Logarithm)

Variable	Constant	Teacher/pupil ratio	No free lunch	Attendance	R ²	n
Reading	-17.26	-0.30 (-1.70)	0.23 (4.98)	4.26 (3.93)	.58	97
Mathematics	-9.51	-0.21 (-1.55)	0.16 (4.39)	2.72 (3.23)	.51	97
Combined	-12.83	-0.24 (-1.67)	0.19 (4.95)	3.38 (3.76)	.57	97

Note. Dependent variable: Natural logarithm of the percentage of students in the school above the national median score. The t ratios are reported in parentheses. The t ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36.

Table 7.—Regression Results: Professional/Pupil Ratio Used as School Input (All Variables Were Converted to the Natural Logarithm)

Variable	Constant	Professional/pupil ratio	No free lunch	Attendance	R ²	n
Reading	-16.59	-0.10 (-0.67)	0.27 (6.47)	4.25 (3.76)	.57	97
Mathematics	-9.31	-0.03	0.19 (5.75)	2.79 (3.18)	.49	97
Combined	-12.4	-0.06 (-0.45)	0.22 (6.43)	3.42 (3.64)	.56	97

Note. Dependent variable: Natural logarithm of the percentage of students in the school above the national median score. The *t* ratios are reported in parentheses. The *t* ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36.

Table 8.—Regression Results: Expenditure per Pupil Used as School Input (All Variables Were Converted to the Natural Logarithm)

Variable	Constant	Expenditure per pupil	No free lunch	Attendance	R ²	n
Reading	-14.58	-0.14 (-0.86)	0.26 (5.76)	4.12 (3.58)	.57	97
Mathematics	-9.08	-0.02 (-0.16)	0.18 (5.31)	2.80 (3.12)	.49	97
Combined	-11.42	-0.07 (-0.54)	0.22 (5.82)	3.36 (3.51)	.56	97

Note. Dependent variable: Natural logarithm of the percentage of students in the school above the national median score. The *t* ratios are reported in parentheses. The *t* ratios for significance levels of 10, 5, and 1% for a one-tailed test were 1.29, 1.66, and 2.36.

student performance. The importance of student socioeconomic status is consistent with other studies. Also consistent with other studies is the finding that school input measures do not exhibit a significant positive influence on student performance.

Studies of this nature continue to be performed, in part because of the lack of a unanimity concerning the influence of various factors on student performance. Moreover, there appears to be heightened interest in and contentiousness of studies of student performance. This can be partially attributed to concerns regarding the declining performance of American students (Chubb & Moe, 1990; Hanushek, Rivkin, & Jamison, 1992). Also, the issue of school finance reform and legal challenges to the method by which public schools are financed promises to increasingly place the questions regarding differences in school resources and student performance addressed in this literature in the legal and public policy spotlight (Dayton, 1993). Furthermore, recent major federal legislation in the form of the Goals 2000: Educate America Act (1994) focuses on, among other things, student performance and the ways in which school resources affect it.

What do the findings here regarding the attendance-performance relationship imply for education analysts and policy makers? One could draw the conclusion that there is

prima facie evidence to support the view that devoting resources to increasing attendance rates is warranted. This conclusion, however, presumes that policies and programs could in fact increase attendance rates and, moreover, do so in a manner that is cost effective. Future analysts should examine both the ability to influence attendance and the cost of such policies and programs. The extent to which latent variables that affect student performance (e.g., student motivation, parental concern) are embedded in the attendance measure, although a difficult empirical matter, also warrants further examination. Attendance per se may not have as strong an effect on performance as implied here because of such latent variable measurement.

The results here are based on data that were aggregated to the school level, and the results should accordingly be interpreted at that level. One cannot presume that student level or classroom level analyses would reveal such strong results. Aggregated data will have less variability than the underlying data measured at the student or classroom levels. One would expect stronger results (i.e., higher *R*-squared values, increased precision of regression coefficient estimates), with aggregated data. Analysts with student- or classroom-level data available for all of the variables (unlike this data set) may find that a multilevel modeling approach (Seltzer, 1994) provides fertile ground for future

research. The influence of attendance on student performance may or may not differ substantially by school or teacher. This is an important and potentially illuminating issue to address. In any case, analysts are well advised to further document the robustness of the attendance–performance relationship with other data because the findings reported here should not be viewed as the last word on the matter.

NOTES

A Summer Faculty Fellowship from the University of Maryland Baltimore Graduate School supported this project. I thank Virginia McConnell, Dave Greenberg, and the anonymous referees for their useful comments. I also thank participants at the 1993 Atlantic Economic Conference in Philadelphia for their comments on an earlier version of this article.

1. Regression of no free lunch on attendance yielded an R-squared value of .29. Regression of no free lunch and percentage minority on attendance yielded an R-squared value of .36. Thus, 71% and 64% variance in attendance were unexplained by the socioeconomic measure or measures; this unexplained variation in the attendance variable must be explained by other factors. These results suggest that attendance and socioeconomic status were not tightly linked.

2. A similar but not identical unit-free measure is the standardized regression coefficient. It is the regression coefficient times the ratio of the standard deviation of the respective independent variable to the standard deviation of the dependent variable. In the context here, the output elasticity appears to have a more meaningful interpretation. See Pindyck and Rubinfeld (1991, pp. 85–86) for a further discussion of these two measures.

3. The output elasticity $(Q/X)/(Q/X)$ will be the same regardless of the level of the dependent or independent variables. For this reason, this model is sometimes referred to as the *constant elasticity* model.

4. The standard scaled condition index values are not reported in these tables because, as Belsley (1991, p. 281) discussed, these are “distorted diagnostic magnitudes” for logarithmically transformed variables.

REFERENCES

- Baltimore, Maryland State Department of Education. *Enrollment by race/ethnic categories*. Baltimore: Author.
- Belsley, D. A. (1991). *Conditioning diagnostics: Collinearity and weak data in regression*. New York: Wiley.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.
- Berger, M. C., & Toma, E. F. (1994). Variation in state education policies and effects on student performance. *Journal of Policy Analysis and Management*, 13(3), 477–491.
- Brown, B., & Saks, D. (1975). The production and distribution of cognitive skills within schools. *Journal of Political Economy*, 83(3), 571–593.
- Caldas, S. J. (1993). Reexamination of input and process factor effects in public school achievement. *The Journal of Educational Research*, 86(4), 206–214.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools*. Washington, DC: The Brookings Institution.
- Citizens Planning and Housing Association. (1990). *Report card on Baltimore City public elementary schools*. Baltimore, MD: Author.
- Dayton, J. (1993). Correlating expenditures and educational opportunity in school funding litigation: The judicial perspective. *Journal of Education Finance*, 19, 167–182.
- Eberts, R. W., Schwartz, E. K., & Stone, J. A. (1990). School reform, school size, and student achievement. *Economic Review, Federal Reserve Bank of Cleveland*, 26(2), 2–15.
- Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, 28(2), 465–498.
- Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis*, 13(2), 189–202.
- Goals 2000: Educate America Act of 1994, Pub. L. No. 103–227, 108 Stat. 125 (1994).
- Gujarati, D. N. (1988). *Basic econometrics* (2nd ed.). New York: McGraw-Hill.
- Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, 14(3), 351–381.
- Hanushek, E. A. (1981). Throwing money at schools. *Journal of Policy Analysis and Management*, 1(1), 19–41.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141–1177.
- Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, 18(4), 45–51.
- Hanushek, E. A. (1991). When school finance reform may not be good policy. *Harvard Journal on Legislation*, 28(2), 424–456.
- Hanushek, E. A. (1994). Money might matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher*, 23(4), 5–8.
- Hanushek, E. A., Rivkin, S. G., & Jamison, D. T. (1992). Improving educational outcomes while controlling costs. *Carnegie-Rochester Conference Series on Public Policy*, 37, 205–238.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994a). Does money matter? Meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5–14.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994b). Money does matter somewhere: A reply to Hanushek. *Educational Researcher*, 23(4), 9–10.
- Johnston, J. (1984). *Econometrics methods* (3rd ed.). New York: McGraw-Hill.
- Kennedy, P. (1992). *A guide to econometrics* (3rd ed.). Cambridge, MA: The MIT Press.
- Murnane, R. J. (1991). Interpreting the evidence on “Does money matter?” *Harvard Journal on Legislation*, 28(2), 457–464.
- Oates, W. E. (1977). The use of local zoning ordinances to regulate population flows and the quality of local services. In O. C. Ashenfelter and W. E. Oates (Eds.), *Essays in labor market analysis* (pp. 201–219). New York: Wiley.
- Pindyck, R. S., & Rubinfeld, D. L. (1991). *Econometrics models and economic forecasts* (3rd ed.). New York: McGraw-Hill.
- Romer, D. (1993). Do students go to class? Should they? *Journal of Economic Perspectives*, 7(3), 167–174.
- Salvatore, D. (1993). *Managerial economics* (2nd ed.). New York: McGraw-Hill.
- Sander, W. (1993). Expenditures and student achievement in Illinois. *Journal of Public Economics*, 52(3), 403–416.
- Seltzer, M. H. (1994). Studying variation in program success: A multilevel modeling approach. *Evaluation Review*, 18(3), 342–361.
- U.S. Department of Education, National Center for Education Statistics. (1990). *Digest of education statistics*. Washington, DC: Author.
- Walberg, H. J., & Walberg, H. J., III (1994). Losing local control. *Educational Researcher*, 23(5), 19–25.
- Wendling, W., & Cohen, J. (1981). Education resources and student achievement: Good news for schools. *Journal of Education Finance*, 7, 44–65.