Economics of Education Review, Vol. 5, No. 1, pp. 41–48, 1986.
Printed in Great Britain.

0272–7757/86 $3.00 + 0.00

Some New Evidence on Teacher and Student Competencies

ROBERT P. STRAUSS* and ELIZABETH A. SAWYER†

*Center for Public Financial Management, School of Urban and Public Affairs, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.

and

†Department of Economics, The Pennsylvania State University, University Park, PA 16801, U.S.A.

Abstract — This paper reports statistical analysis of the determinants of average student performance on standardized examinations, and also the determinants of the extent to which students fail such examinations. Unlike most other cross-sectional studies of performance among school districts within one state, this study uses the quality of teachers, as measured by standardized test scores, as a determinant of performance. Perhaps the most striking empirical result of the study is the finding that a 1% increase in teacher quality, as measured by standardized test scores, is accompanied by a 5% decline in the rate of failure of students on standardized competency examinations. The corresponding impact on average or mean achievement of teacher quality is, by contrast, quite modest: 0.5–0.8% per 1% improvement in teacher quality.

INTRODUCTION

SINCE the publication of the Coleman Report (1966) there have been numerous statistical investigations of student achievement in relation to school inputs, socio-economic background and motivational factors. Hanushek (1979, 1981) provides excellent critical reviews of this educational production function literature. Until the recent papers of Boardman *et al.* (1971, 1973, 1976) and Summers and Wolfe (1977), there had been relatively little evidence brought to bear on the conjecture that varying school inputs would materially alter student achievement. The absence of strong input–output or production relationships has been attributed to failing to properly identify the output(s) of interest, as well as failing to identify the underlying structural (as opposed to reduced form) relationships between factors of production and outputs. The recent findings of significant school input–output relations are interesting, because the models go well beyond economic production analysis and appeal to a broader and more numerous set of inputs.[1]

Our study differs from these in that sizeable and significant production relations are found in the context of a straightforward Cobb–Douglas production function. We find that when we focus on the failures of the educational process, we can explain better than 80% of the variation in the absolute numbers of students who fail high school competency exams in the spring of their junior year. This level of explanation is obtained with only six basic inputs which correspond to the economic notions of capital and labor intensity. These basic inputs similarly explain a sizeable fraction of the variation in the rate of failure of students who took the high school competency exams.

Perhaps our most startling finding is that a 1% increase in teacher quality, *ceteris paribus*, as measured by standardized test scores, is accompanied by a 5% decline in the level of failure or rate of failure of students on high school competency

[Manuscript received 24 April 1985; revision accepted for publication 3 July 1985.]

examinations. The corresponding partial elasticity for teacher quality *vis-à-vis* mean student achievement is much more modest (0.5–0.8%).

A MODEL OF EDUCATIONAL PRODUCTION

Much of the difficulty in isolating significant production relations beyond those difficulties noted by Hanushek (1979, 1981) is due to the failure to distinguish between inputs to the production process and factors contributing to the size of the budgets which school administrators then allocate. For example, it has been common[2] to write an achievement function, relating student achievement at the student or school district level, i as:

$$\text{achievement} = f(\text{expenditures/pupil}_i, \text{pupil/teacher}_i, \text{building area}_i, \text{average family income}_i, \text{etc.}).$$

Note that in the above specification that total resources per pupil is a regressor along with average family income. Now, parental or average family income per district, used in the above achievement function, probably also drives expenditures in the ordinary demand-for-public-goods sense which determines the size of the budget. Thus, to the extent that expenditures are really endogenous to income, it is likely that the two are highly correlated, and the resulting single equation regression estimates are likely to be misleading.

Another difficulty with the above sort of specification is that total dollar expenditures equal total factor payments for inputs. If factor prices are constant across observations, which is reasonable in a cross-section context, then input quantities will be monotonic with total expenditures. The intercorrelation again could well cause various factors of production then to be insignificant in the regression which seeks to explain student achievement.

Thus, in our view an achievement function such as the above contains both a demand-for-public-goods relation (with income and total expenditures on the right-hand side, and input quantities which are monotonic with expenditures) and a production relation. It is not surprising that school inputs have generally been either small in impact or statistically insignificant, because of the obvious collinearity between expenditures and various factors of production on the right hand side. In the extensive public finance literature,[3] expenditures would be endogenous and income a causal factor in the expenditure equation. If one believes that school administrators allocate a fixed budget — itself the result of a public choice process — with an objective of maximizing passes or its dual, in our view, minimizing failures, then a good many regressors which typically are used in educational production function studies disappear since they determine expenditure levels rather than achievement *per se*.

In our view, then, achievement depends simply on the factors available to administrators to produce education: students, teachers, capital and some measure of motivation. Most socio-economic factors used in earlier educational production studies can be viewed as driving the expenditure equation rather than the achievement equation since attitudes and ability to pay will affect the size of the budget, which is then allocated to educational production. It is possible to include race in the production relation as well as the expenditure relation; however, race-effects in the production relation will reflect both the effects of previous discrimination in the provision of educational services and other possible effects.

The important question remaining for our analysis is the specification of the proper unit of output. Most studies of educational achievement have focused on mean or median achievement. Our empirical analysis is in that tradition; however, we shall also estimate the impact of factors of educational production on those most at risk in the system by examining how factors of production affect the number who fail achievement tests.

In the context of cross-sectional estimation, we take budget size to be exogenous, and thus estimate only the production function. Our cross-section Cobb–Douglas production relationship with the usual assumptions of constant, unitary elasticity of substitution among factors then is:

$$\text{failure}_{ij} = \beta_0 * \text{class}_{ij}^{\beta_1} * \text{NTE}_{ij}^{\beta_2} * \text{assets}_{ij}^{\beta_3} * \text{college}_{ij}^{\beta_4} * \text{blacks}_{ij}^{\beta_5} * \text{teachers}_{ij}^{\beta_6} * \epsilon_{ij}. \quad (1)$$

For the ith school district and jth grade, failure is the number of students in the class who fail a competency exam; class measures the number of students in the ith district and jth grade; NTE is a measure of average teacher quality;[4] assets is a measure of the capital value in the school district; college, a proxy for student motivation, is the number of students indicating an interest in attending either a community college, technical institute or private college or university; blacks is the number of black students; and teachers is the number of teachers.

We expect β_1 and $\beta_5 > 0$ and $\beta_2 \ldots \beta_4$ and $\beta_6 < 0$.

Since NTE is an average value for an entire district, it may be more natural to write a relative production relationship which determines the rate of failure:

$$\left(\frac{\text{failure}}{\text{class}}\right)_{ij} = \theta_0 * \text{NTE}_{ij}^{\theta_1} * \left(\frac{\text{assets}}{\text{class}}\right)_{ij}^{\theta_2} * \left(\frac{\text{college}}{\text{class}}\right)_{ij}^{\theta_3} * \left(\frac{\text{blacks}}{\text{class}}\right)_{ij}^{\theta_4} * \left(\frac{\text{class}}{\text{teachers}}\right)_{ij}^{\theta_5} * \mu_{ij}. \quad (2)$$

It will be of interest to compare the results of estimating variants of (1) and (2) with a more traditional specification in which the mean achievement score of an achievement test is the dependent variable of interest:

$$\text{average achievement}_{ij} = \gamma_0 * \text{NTE}_{ij}^{\gamma_1} * \left(\frac{\text{assets}}{\text{class}}\right)_{ij}^{\gamma_2} * \left(\frac{\text{college}}{\text{class}}\right)_{ij}^{\gamma_3} * \left(\frac{\text{blacks}}{\text{class}}\right)_{ij}^{\gamma_4} * \left(\frac{\text{class}}{\text{teachers}}\right)_{ij}^{\gamma_5} * \eta_{ij}. \quad (3)$$

With regard to (2), we expect $\theta_1 \ldots \theta_3 < 0$, and θ_4 and $\theta_5 > 0$. With regard to (3), we also expect $\gamma_1 \ldots \gamma_3 > 0$, and γ_4 and $\gamma_5 < 0$. That is, we generally expect more and higher quality resources, and more highly motivated students to improve performance, and entertain the possibility that black students may perform more poorly.

To estimate (1)–(3), data were collected on the 145 school districts in North Carolina for 1977–1978. The historical records of the State Department of Education provided reliable data on the number of teachers, full-time equivalent student enrollees, those high school students interested in post-high school education, the racial composition of the schools, the insured value of the capital stock in the school district for 105 districts, and average National Teacher Evaluation score given by the Educational Testing Service of new teachers in each district.

Two measures of failure were available: (1) the fraction of high school juniors failing the reading and mathematics competency examinations in the Spring of their junior year which were developed by the State in conjunction with educational researchers at the University of North Carolina and, after a pretest, administered in Spring 1978; and (2) the average achievement on the Norm Referenced Achievement Test which was administered to the same high school juniors in 1978.

Before turning to the empirical results, several observations about equations (1)–(3) are in order. First, the production function is cross-sectional in nature at $j = 11$. It would be most desirable to have data on a number of schools over a period of time so that one could observe actual variations in factor intensity per school district in relation to changes in levels of output. Unfortunately this is rarely the case in most traditional production studies, and such data are unavailable to us here. Second, while (2) attempts to state the production relation in terms of rates, there are some non-comparabilities in the measures being used. For example, the dependent variable is defined in terms of the number of students failing the high school equivalency examination in comparison to the numbers in the class taking the examination. However, the student–teacher ratio measures the total number of teachers in the school district divided by the total number of students enrolled in the school district. Thus the class-size measure is an aggregate measure across all levels or grades, and does not pertain to just those in their final grade of high school. Similarly, the measure of non-white enrollment and the measure of capital intensity both refer to total figures in each district. This may not reflect the resources and environment facing students taking the competency examination. On the other hand, to the extent that failure and achievement reflect the accumulated impact of school resources, it is appropriate to measure inputs on average across grades, rather than at the margin. However, one still must assume that observed pupil–teacher ratios in 1978 are representative of what students historically experienced.

Third, the measure of teacher quality refers to the average NTE test score of teachers in a school district, and thus does not correspond directly to the quality of teachers who instructed the students over time. Another difficulty with this specification is that it involves district-by-district analysis, rather than school-by-school or pupil-by-pupil analysis.

While we are clear about the shortcomings of the operational measures used to estimate the production relationship, we do note that our analysis is performed within the confines of one state and the data do reflect the experience of entire school districts. Thus the results for the state may be viewed as representative of the state as well as being institutionally consistent. Also, the types of data being employed are broadly similar to data used in earlier studies which found little relationship between school inputs and outputs.

Table 1 displays the means and standard deviations of the data. On average 10% of the students failed the reading portion of the competency examination while 14.8% failed the mathematics portion of the competency examination. The average district had 500 students in the senior class, of whom 31% were black; the average pupil–teacher ratio was 19.

The estimation results of double-log versions of (1)–(3) are shown in Tables 2 and 3. Since capital data were available for only 105 of the 145 districts, estimates were made with and without the capital variable in the model. Panel A of Table 2 shows the estimation results for (1). Better than 83% of the variation in the total number of failures is explained by the six inputs. By far the most significant and sizeable variable which determines failures is the average teacher quality in the school district. The elasticity is always above −5.0 when the capital variable is in, and closer to −6.0 when the capital variable is not in.[5] Surprisingly, class size, number of teachers and number with post-high school educational intentions are not stastically different from zero.[6] The race measure is statistically significant in 2/4 of the regressions in panel B, and suggests that the more blacks in a school district, the more who will fail. Note, however, that the size of this effect is quite small: no elasticity exceeds 0.10.

The results for (2) are in panel B of Table 2. Generally, the level of explained variance is about half that of panel A. Again, the quality of the teachers as measured by their test scores has highly significant and very large effects on the rate of reading and math failures. Again, the presence of black students increases the failure rate in a significant but very modest fashion. The impact of teacher quality on the rate of failure varies from an elasticity of 5.02 to one of 6.30, while the elasticity of the effect of race on the rate of failure varies between 0.060 and 0.089.

The estimation results for the average performance in each district are shown in Table 3. As in panel B of Table 2, the level of explained variance is relatively strong — between 33 and 60%; however, there the similarities end, for the impact of the various inputs on average achievement is quite

Table 1. Means and standard deviations of data*

	μ	σ
Reading failure on competency exam	54.069	68.23
Math failures on competency exam	79.145	98.67
Class (j = 11)	554.617	681.13
NTE	1172.510	41.96
Assets	$11,320,510	$16,515,798
College	173.097	272.82
Black	165.872	233.42
Teachers	28.548	36.27
Reading failure rate	0.1038	0.048
Math failure rate	0.1479	0.062
Assets/class	$2169	$484
College/class	0.2937	0.0905
Black/class	0.3132	0.2111
Class/teachers	19.435	1.36
Reading achievement	105.064	3.39
Math achievement	96.668	5.77

*Figures refer to 145 observations except for asset data. See text for definitions.

Table 2. OLS estimates of model (1) and (2) (t ratio in parentheses)

Panel A: model (1)

Dependent variable†	Log B_0	B_1 log(class)	B_2 log(NTE)	B_3 log(assets)	B_4 log(college)	B_5 log(blacks)	B_6 log(teachers)	$\bar{R}^2$	n
1. Log (reading failures)	42.5778 (4.78)	0.3306 (0.57)	−5.8759 (−4.53)	−0.1165 (−0.63)	−0.1401 (−1.01)	0.0913 (2.57)	0.8042 (1.22)	0.8463	105
2. Log (reading failures)	42.9371 (5.42)	0.6489 (1.44)	−6.2374 (−5.42)	—	−0.1560 (−1.16)	0.0930 (2.58)	0.3941 (0.80)	0.8386	145
3. Log (math failures)*	38.5083 (4.49)	0.4564 (1.83)	−5.1138 (−4.11)	−0.2502 (−1.41)	−0.1702 (−1.29)	0.0553 (1.62)	0.9447 (1.49)	0.8638	105
4. Log (math failures)	40.9523 (5.43)	0.5611 (1.23)	−5.8742 (−5.65)	—	−0.2065 (−1.82)	0.0595 (1.92)	0.6274 (1.24)	0.8593	145

Panel B: model (2)

Dependent variable†	Log A_0	A_1 log(NTE)	A_2 log(assets/class)	A_3 log(college/class)	A_4 log(blacks/class)	A_5 log(class/teachers)	$\bar{R}^2$	n
5. Log (reading failure/class)*	43.3710 (5.24)	−5.9533 (−4.64)	−0.1080 (−0.62)	−0.1541 (−0.96)	0.0854 (2.21)	−0.9793 (−1.62)	0.4655	105
6. Log (reading failure/class)	43.6891 (5.63)	−6.3253 (−5.88)	—	−0.1645 (−1.40)	0.0891 (2.83)	−0.4841 (−0.98)	0.4173	145
7. Log (math failure/class)	37.5729 (4.45)	−5.0228 (−4.06)	−0.2602 (−1.47)	−0.1537 (−1.18)	0.0623 (1.91)	−0.7378 (−1.32)	0.4078	105
8. Log (math failure/class)	39.4046 (5.32)	−5.6943 (−5.55)	—	−0.1890 (−1.68)	0.0674 (2.25)	−0.4393 (−0.93)	0.3726	145

*Standard errors of estimate have been corrected as suggested by White (1980) after heteroskedasticity was found at the 1% level.
† All variables stated in natural logarithms. See text for variable definitions and data sources.

Table 3. OLS Estimates of model (3) (t ratio in parentheses)

Dependent variable†	Log c_0	c_1 log(NTE)	c_2 log(assets/class)	c_3 log(college/class)	c_4 log(blacks/class)	c_5 log(class teachers)	$\bar{R}^2$	n
1. Log (average reading score	0.5694 (1.11)	0.5225 (6.82)	0.01448 (1.27)	0.0126 (1.38)	−0.0022 (−0.99)	0.0987 (2.80)	0.6058	105
2. Log (average reading score)	0.3999 (0.38)	0.5773 (9.27)	—	0.0166 (2.44)	−0.0025 (−1.34)	0.0647 (2.25)	0.5579	145
3. Log (average math score)	−1.2854 (−0.93)	0.7132 (3.52)	0.0301 (1.04)	0.0316 (1.49)	−0.0053 (−0.99)	0.2066 (2.25)	0.3281	105
4. Log (average math score)	−1.5260 (−1.36)	0.8064 (5.20)	—	0.0379 (2.23)	−0.0061 (−1.33)	0.1468 (2.06)	0.3342	145

* Standard errors of estimate have been corrected as suggested by White (1980) after heteroskedasticity was found at the 1% level.
† All variables stated in natural logarithms. See text for variable definitions and data sources.

Table 4. OLS Estimates of model 2 and 3 with per capita income (t ratio in parentheses)

Dependent variable*	Log F_0	F_1 log NTE	F_2 log (assets/class)	F_3 log (college/class)	F_4 log (blacks/class)	F_5 log(class/ teachers)	F_6 log(per capita income)	$\bar{R}^2$	n	f
2.5. Log (reading failures/class)	43.3710 (5.24)	−5.9533 (−4.64)	−0.1080 (−0.62)	−0.1541 (−0.96)	0.0854 (2.21)	−0.9793 (−1.62)	—	0.4655	105	17.24
Log (reading failures/class)	34.9092 (4.01)	−3.4373 (−2.41)	−0.0985 (−0.56)	−0.1475 (−1.15)	0.1511 (4.02)	−0.8613 (−1.55)	−1.1476 (−3.40)	0.5217	105	17.81
2.7. Log (math failures/class)	37.5729 (3.67)	−5.0228 (−2.30)	−0.2602 (−1.46)	−0.1537 (−1.18)	0.0623 (1.91)	−0.7378 (−1.32)	—	0.4078	105	13.63
Log (math failures/class)	31.6054 (3.67)	−3.2484 (−2.30)	−0.2535 (−1.46)	−0.1489 (−1.18)	0.1085 (2.92)	−0.6545 (−1.19)	−0.8094 (−2.42)	0.4412	105	12.90
4.1. Log (average reading score)	0.5694 (1.11)	0.5225 (6.82)	0.0145 (1.27)	0.0126 (1.38)	−0.0022 (−0.99)	0.0987 (2.80)	—	0.6058	105	30.43
Log (average reading score)	1.0976 (2.29)	0.3655 (4.66)	0.0139 (1.45)	0.0122 (1.73)	−0.0063 (−3.03)	0.0914 (3.00)	0.0716 (3.85)	0.6577	105	31.39
4.3. Log (average math score)	−1.2854 (−0.93)	0.7132 (3.52)	0.0301 (1.04)	0.0316 (1.49)	−0.0053 (−0.99)	0.2066 (2.25)	—	0.3281	105	9.67
Log (average math score)	−0.8091 (−0.56)	0.5716 (2.42)	0.0296 (1.02)	0.0312 (1.47)	−0.0090 (−1.44)	0.1999 (2.18)	0.0646 (1.16)	0.3371	105	8.31

* All variables stated in natural logarithms. See text for variable definitions and data sources.

different than on the rate of failure. First, teacher quality has only a modest effect on achievement: the elasticity of teacher quality with respect to reading and mathematics scores varies between 0.5 and 0.8. Second, the fraction of black students does not have a statistically significant effect on achievement, whereas it does on the level and rate of failure. Third, the pupil–teacher ratio does significantly affect average achievement while it does not significantly affect the rate of failure, and in a surprisingly positive manner. Apparently, larger class size tends to lead to improved average performance rather than to lower average performance; however, the effect is quite modest, as the elasticity of the pupil–teacher ratio with respect to average test scores varies from 0.09 to 0.20. Finally, the measure of motivation shows some modest impact on average achievement which it does not show on the rate of failure.

Since earlier studies use income as a direct, explanatory variable in achievement regression equations, we estimated variants for (2) and (3) which include the log of *per capita* income on the right-hand side. The estimation results (with the comparison equations from Tables 2 and 3) are in Table 4. Several observations are immediately in order: first, the effects of assets, interest in higher education, and pupil–teacher ratio are stable in impact between the two specifications both in terms of the size and reliability of effects. Second, per capita income *per se* operates in a significant fashion with regard to the failure rate in reading (an elasticity of −1.1) and with regard to the failure rate in mathematics (an elasticity of −0.81). Further, higher income districts tend to have lower failure rates.

The effects of teacher quality and race change to a greater degree than the other inputs when *per capita* income is in the regression equation. The race elasticities double in size but remain in absolute magnitude rather small (none is over 0.15), while the impact of teacher quality falls by as much as 42%. Thus, with per capita income as a regressor, the elasticity of failure with regard to teacher quality drops from a range of 5.02–5.95 to 3.25–3.44 but the effects of teacher quality are still very large when *per capita* income is added as an additional regressor.

Of correlative interest, and consistent with our view that income is a demand for public goods characteristic, is the result that each of these resource and attitude measures' impact is smaller with per capita income in the regression equation than without it. Income is a correlate for resource availability and also picks up differences in student outlook.

DISCUSSION

The estimation results contained in Tables 1 and 2 contain a fairly dramatic and, at least to these authors, plausible set of results. First, teacher quality makes an enormous difference in affecting whether or not sizeable numbers of students fail to demonstrate independently measured reading and mathematics skills. Second, teacher quality also makes a difference with regard to average performance in the classroom; however, the size of the effect is one-tenth that of the impact of teacher quality on failure levels or rates. Third, the pupil–teacher ratio does not affect the level or rate of failure but does impact modestly in a surprisingly positive fashion on average achievement. Fourth, the racial composition of school districts does affect the extent of failure, but not average achievement (though with one exception). This first effect of race may well reflect the accumulated impact of previous discrimination in the equal provision of educational services within a district, and the lingering effects of a previously segregated school system. Accounting for the income in each school district did not change the character of these results, although teacher quality's impact drops from a five-fold elasticity to a three-fold elasticity. Finally, we observe that analysis at the school district level can provide important insights, and that differentiating between average achievement and the performance of those educationally at risk yields important insights.[7]

Of the inputs which are potentially policy-controllable (teacher quality, teacher numbers via the pupil–teacher ratio and capital stock) our analysis indicates quite clearly that improving the quality of teachers in the classroom will do more for students who are most educationally at risk, those prone to fail, than reducing the class size or improving the capital stock by any reasonable margin which would be available to policy-makers. The size of this differential impact among inputs is enormous, and undoubtedly deserves further examination; however, at this juncture in our research, it is unmistakably clear that teachers matter far more than has been previously documented by other researchers in the field.

Acknowledgements — The authors wish to thank the North Carolina Department of Education for providing unpublished data, and Denise DiPasquale, Steve Garber, Knox Lovell, Peter Schmidt and the Editor for valuable comments. Responsibility for remaining errors rests with the authors.

NOTES

1. The initial Boardman *et al.* (1971, 1973, 1976) model begins with 6 endogenous variables and 41 inputs of various types, while Summers and Wolf (1977) experiment with 29 different inputs.
2. See, for example, various studies reviewed by Guthrie (1970) which relate expenditures to achievement.
3. See, for example, Gramlich (1972) or Inman (1979).
4. In particular, NTE is the average composite score reported by teachers in the districts under study on the National Teacher Exam.
5. Summers and Wolfe (1977) found that students with teachers from better schools did better, but that higher NTE scores were associated with lower test scores. On the other hand, our results for capital correspond to theirs.
6. Since the average class size is 19.4, it is possible that the positive relationship between class size and average achievement is capturing possible economies of scope within the observed range of variation in class size. Recall that the standard deviation in the average class size was only 1.3 students per teacher. To the extent that somewhat larger schools are more successful, the class size measure as utilized may be capturing this effect as well.
7. Note also that our sign expectations were confirmed in each instance for 62 of the 66 parameters estimated.

REFERENCES

Boardman, A.E., Davis, O.A. and Sanday, P.R. (1971) A simultaneous equations model of the educational process. *J. publ. Econ.* 7, 23–50.

Boardman, A.E., Davis, O.A. and Sanday, P.R. (1973) A simultaneous equations model of the educational process; the Coleman data revisited with an emphasis upon achievement, pp. 62–71. Proceedings of the Social Statistics Section, American Statistical Association.

Boardman, A.E., Sanday, P.R., Davis, O.A. and Johnson, N. (1976) Simultaneous equations models of the educational process. In *The Evaluation of Social Programs* (Edited by Abt, C.C.) pp. 375–406. Beverly Hills, CA: Sage.

Coleman, J.S. (1966) *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

Gramlich, E.M. (1972) Intergovernmental grants: a review of the empirical literature. In *The Political Economy of Fiscal Federalism* (Edited by Oates, W.). Lexington, MA: D.C. Heath & Co.

Guthrie, J.W. (1970) A survey of school effectiveness studies. In *Do Teachers Make a Difference?* pp. 25–54. Washington, DC: U.S. Government Printing Office.

Hanushek, E.A. (1979) Conceptual and empirical issues in the estimation of educational production functions. *J. hum. Resources* **XIV**, 351–388.

Hanushek, E.A. (1981) Throwing money at schools. *J. Policy Anal. Manage.* **1**, 19–42, 53–55.

Inman, R. (1979) Fiscal performance of local governments: an interpretative review. In *Current Issues in Urban Economics* (Edited by Mieszkowski, P. and Straszheim, R.). Baltimore, MD: Johns Hopkins Press.

Summers, A.A. and Wolfe, B.L. (1977) Do schools make a difference? *Am. econ. Rev.* **67**, 639–652.

White, H. (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.