Functional form and the estimated effects of school resources

David N. Figlio *

Department of Economics, University of Florida, Gainesville, FL 32611-7140, USA

Received 19 February 1998; accepted 1 August 1998

Abstract

A widely-cited result in the education production function literature is that the level of measured school inputs (e.g., student–teacher ratio or starting teacher salary) is not associated with differences in student achievement, all else equal. I argue that this result may be attributable in part to the restrictive functional form assumptions used in the existing literature. I use detailed national student-level data to estimate an education production function that does not employ the restrictive assumptions of homotheticity and additivity and find statistically significant evidence that school inputs are associated with student performance. Moreover, I find evidence that traditional education production functions may understate the magnitude of the effects of school factors on student achievement. However, I find that even these larger estimated effects of measured school inputs are very small. [JEL I21] © 1999 Published by Elsevier Science Ltd. All rights reserved.

Keywords: Functional form; School quality; Student test scores

1. Introduction

Significant differences in school quality exist in the United States, and few would argue that schools do not contribute to student learning. However, there exists considerable controversy as to whether *measured school inputs* matter in determining student achievement. Much of this controversy stems from the observation that while real per pupil expenditures on education have increased almost monotonically over the last 30 years, average student performance on standardized tests has stagnated.

Since the mid-1960s, scores of scholarly research articles and books have investigated the relationship between expenditures (or measured school inputs) and student performance, and a majority of these find either no statistically significant relationship between school inputs and student performance, or find significant coefficient estimates with the “wrong” sign (e.g., higher student–teacher ratios associated with higher student test

scores.) In a widely-cited article, Hanushek (1986) summarizes the existing literature by reporting that “there is no systematic relationship between school expenditures and student performance,” a finding echoed in his more recent summaries. Furthermore, Hanushek, Rivkin, & Taylor (1996) argue that many of the studies that find positive and significant relationships between school inputs and student performance use aggregate data that will bias the coefficient estimates upward. They show that when only studies employing individual-level data are considered, the evidence supporting a link between measured school inputs and student achievement becomes even more tenuous.

While the evidence supporting the hypothesis that changes in school inputs are not related to changes in student outcomes is plentiful, the results of the existing education production literature are not fully convincing. Despite the widespread academic evidence that suggests no link between school inputs and student achievement, there is a general popular belief that more school inputs are associated with better performance. For example, the best educated, most sophisticated parents tend to settle

* E-mail: Figliodn@dale.cba.ufl.edu

in school districts that tax highly to spend more on schools. If well-educated people believed that this additional spending was not useful, it is doubtful that they would choose to continue to spend as much on schooling. Furthermore, many parents donate money and time to increase school resources, precisely because they believe that school inputs matter. So-called “school success stories” tend to have several common factors—smaller classes, school-based management, a good principal, and more money. And California recently passed an initiative to substantially reduce public school class sizes in early grades. All of these examples suggest a belief that spending more to increase school inputs can result in higher levels of student achievement. In addition, recent work on the effects of fiscal constraints (e.g., Figlio, 1997a; Downes & Figlio, 1998) show that tax limits significantly reduce student outcomes.

There are several possible reasons why education production function researchers and widespread public opinion might arrive at different results. One reason might be that the school inputs used in education production function studies may not be good indicators of what goes on in the classroom. For example, there are highly-paid teachers who perform less well than do many low-paid teachers. In addition, there may be many small classes where the students fare worse than those in much larger classes taught in an innovative way.

There are other reasons in addition to the one mentioned above that contribute to the difference between the education production findings and popular opinion. I do not attempt to better model the classroom-specific features that may explain a school’s contribution to student performance. Instead, given that the “money does not matter” conventional wisdom has been based primarily on the lack of statistical significance, rather than on the magnitude of relationships between measured school inputs and student achievement (see, e.g., Hanushek, 1986, for an example), I question whether these results could be driven in part by the assumptions made by the researcher. Existing individual-level studies concerned with the determinants of student outcomes make a number of simplifying assumptions. For instance, individual-level studies generally require that the education production relationships that they estimate be additive in the inputs of production.¹ Any systematic differences in the ways in which school inputs affect heterogeneous people are not captured in the estimation.

Another simplifying assumption made in the existing education production literature is that the education production relationship is homothetic—that is, the marginal

rate of substitution among inputs of education production depends only on the proportions of the inputs and not on the scale of production. A homothetic education production function implies that the marginal effects of school services (or parental participation, or any other input) on student performance will be the same regardless of the scale of production. Hence, the ways in which school inputs, say, affect student achievement are restricted to be the same for students with very low input levels (e.g., low-endowed students in impoverished schools) as for those with uniformly high input levels (e.g., high-endowed students attending rich schools). In short, only the scale of education production would be different between these two scenarios. Many studies go even further and assume constant elasticities of substitution and scale, as in the case of the Cobb–Douglas production function. But if the effects of school inputs, say, on student achievement are sensitive to variation in the scale of education production, these assumptions may not be appropriate. One possible consequence of these assumptions may be reduced precision of estimates.

There exists considerable evidence in the economics and public policy literature that the production function for many quasi-public goods (that is, goods that use both public and private inputs in their production) does not exhibit additivity or homotheticity. For instance, Gyapong & Gyimah-Brempong (1988) argue for a more flexible functional form in estimating the production of public safety, and Duncombe (1992) finds that the production of fire protection does not satisfy the restrictions spelled out above. Furthermore, several articles in the economics of education literature using aggregate data (e.g., Jimenez, 1986; Callan & Santerre, 1990; Gyimah-Brempong & Gyapong, 1992) have shown that estimation of education cost functions require a more flexible functional form. Hence, there is reason to believe that education production functions may not be additive or homothetic either.

In this study I estimate an education production function that does not maintain the assumptions of additivity and homotheticity. I use the detailed, restricted-access version of the National Education Longitudinal Survey (NELS), administered nationally to eighth-graders by the U.S. Department of Education in 1988 and again to the students as tenth-graders in 1990. Because the students were tested twice, I can estimate a value-added measure of student achievement. To examine the effect of functional form assumptions in the education production literature, I use the transcendental logarithmic (translog) functional form introduced by Christensen, Jorgenson, & Lau (1971). A benefit of the translog functional form is that when the appropriate restrictions are made, it is possible to directly test the functional form assumptions of the existing education production literature. When I formally test these restrictions, I find that the assumptions of additivity and homotheticity maintained by the

¹ Some studies, such as Summers & Wolfe (1977), relax this assumption by interacting the school input variables with selected demographic characteristics. However, even in these studies, the assumption of additivity is relaxed in very special cases.

existing literature can be rejected at any reasonable level of significance. While I hesitate to draw conclusions on the basis of specification tests, the results provide some additional evidence that authors in the existing literature may be estimating inappropriate functional forms.

I find evidence of a systematic relationship between measured school inputs and the levels of and changes in student performance. Specifically, I find that the school's student-teacher ratio is negatively related to the student performance, and that the starting teacher salary is positively related to student achievement levels and growth, particularly in schools where student improvement may be most needed. However, I find little evidence that the percentage of teachers with a master's degree or the length of the school year significantly affects student performance. Although I find evidence of a relationship between school inputs and student achievement, the estimated marginal productivities of school inputs seem rather small. Therefore, I argue that if measured school inputs matter in the production of student achievement, money does not seem to matter much.

Furthermore, my findings are highly complementary to those presented in recent work by Eide & Showalter (1998). Eide and Showalter employ quantile regression methods and a different data set to address the question of nonlinear effects of school resources on student performance. While Eide and Showalter and I take different paths, we arrive fundamentally at the same conclusion: traditional specifications of education production functions are overly restrictive because they fail to capture potentially nonlinear effects of school resources.

2. Functional form assumptions in education production relationships

The literature surveyed by Hanushek (1986, 1991) generally assumes that student achievement in time t , y_t , is a function of school inputs S_{t-1} , family inputs F_{t-1} , and peer inputs P_{t-1} . In addition, many authors (e.g., Ehrenberg & Brewer, 1994) argue that what we truly care about when estimating education production relationships is value added, or augmentations to learning. Therefore, these authors estimate a value added, or gain score, equation, in which the set of right-hand-side variables includes y_{t-1} .² Hence, I can write the education production function as

$$y_t = g(S_{t-1}, F_{t-1}, P_{t-1}, y_{t-1}) \quad (1)$$

Most studies in the economics and public policy literature are concerned with measuring $\partial y / \partial S_{t-1}$, and treat the other variables as controls. As the education literature does not provide a priori assistance in determining the appropriate form for $g(\cdot)$, existing individual-level education production function studies assume that $g(\cdot)$ is homothetic. Hence, existing studies assume that the education production relationship is

$$y_t = m(h(S_{t-1}, F_{t-1}, P_{t-1}, y_{t-1})) \quad (2)$$

where $m(\cdot)$ is monotonic and $h(\cdot)$ is homogeneous of degree one. Furthermore, existing studies generally assume that the education production function is additive, so

$$y_t = m_1(S_{t-1}) + m_2(F_{t-1}) + m_3(P_{t-1}) + m_4(y_{t-1}) \quad (3)$$

Christensen, Jorgenson, & Lau (1971) develop a flexible functional form that does not employ additivity and homotheticity as part of the maintained hypothesis. In essence, the translog production function involves geometric expansion of the logarithm of the CES functional form. The resulting functional form provides a local second-order approximation to any production function,³ and in my application can be written

$$\begin{aligned} \ln y_{it} = & \alpha_0 + \sum_k \beta_k (\ln X_{kit-1}) + \frac{1}{2} \sum_k \gamma_{kk} (\ln X_{kit-1})^2 \\ & + \sum_j \sum_k \gamma_{jk} (\ln X_{jit-1}) (\ln X_{kit-1}) \\ & + \sum_k \theta_k (\ln X_{kit-1}) (\ln y_{it-1}) + \alpha_1 (\ln y_{it-1}) \\ & + \frac{1}{2} \alpha_2 (\ln y_{it-1})^2 \end{aligned} \quad (4)$$

for all $j \neq k$. By symmetry, $\gamma_{jk} = \gamma_{kj}$. The vector X_{it-1} includes the school input variables, the peer variables, and the family variables, as described above.

Although Eq. (4), as a linear function, is directly estimable using least squares methods, the most common way to estimate the parameters of the translog function in production applications involves using a system of

² A special case of the value-added specification, used by some researchers, involves estimating an education production function with a dependent variable of $y_t - y_{t-1}$. In this application, I choose to place y_{t-1} as an explanatory variable, rather than expressing the dependent variable in terms of differences, for two reasons: first, including a lagged test score as an explanatory variable increases the flexibility of the model, by not restricting the relationship between past and present test scores to have a coefficient of one. Second, and more practically,

expressing the dependent variable as a difference would have necessitated substantially rescaling the test scores so that I could take logarithms of the dependent variable—a transformation that I would prefer not to make.

³ The translog functional form was also introduced independently by Griliches & Ringstad (1971). These authors were apparently influenced by Hedy & Dillon (1961), who first introduced a geometric expansion similar to the translog production function.

cost-share equations. In such an empirical specification, the researcher uses maximum likelihood to estimate a simultaneous system of equations in which each factor of production's share of total costs of producing the output is a separate dependent variable. On the right-hand-side of each equation would be the prices of the various factors of production. Kim (1992) points out that the cost-share approach is efficiency enhancing over direct estimation because it enables the researcher to utilize more information provided by production theory (e.g., input demand parameters) than directly estimating Eq. (4) would allow.

While such an estimation strategy is sensible when approximating production functions in many economic applications, estimating a system of cost share equations makes less sense with education production functions. The theory of education described above suggests that student learning is derived from many sources—including families and peers—and not just schools. Many of the inputs in the education production function, especially on the individual level, do not have prices *per se*. In addition, it is impossible to measure the total cost of providing a particular individual's education level, so the calculation of cost shares for school or family inputs, say, is untenable. Therefore, any gains in efficiency (in terms of reduced standard errors) from using the system of equations approach would surely be greatly overshadowed by the measurement error that accompanies approximating cost shares and factor prices in an individual-level education production function. Hence, I sacrifice full efficiency to avoid the biases inherent in mismeasuring prices here, and instead estimate Eq. (4) with ordinary least squares, without the use of cost share equations.⁴

I am not particularly interested in the individual estimated parameters, but rather in the production elasticities suggested by the model. Specifically, I am interested in estimating $\partial \ln y / \partial \ln x_{j-1}$, where y is student achievement, holding past achievement constant, and x is the relevant school input (e.g., student-teacher ratio) in the production function. As the form of Eq. (4) suggests, this derivative is a function not only of the direct effect β_j of the input x_j on student performance, but also of the quadratic effect $\gamma_{jj} \cdot x_j$ and of the interactive effects $\gamma_{jk} \cdot x_k$ ($\forall k \neq j$). Therefore, in order to calculate this derivative, one must evaluate $g + 1$ parameters (where g is the number of explanatory variables used in the estimation). I

construct asymptotic standard errors for these elasticities to gauge the statistical significance of the results.⁵

It is evident that the translog education production outlined above does not require additivity or homotheticity; indeed, these assumptions can be directly tested by imposing restrictions on the coefficients of the translog model. Since additivity requires that the production function can be written as $F = F_1(X_1) + F_2(X_2) + F_3(X_3) + \dots$, the interaction terms in the translog model would be constrained to zero if additivity is imposed. Hence, additivity would require that $\gamma_{jk} = 0$ for all $j \neq k$. Furthermore, I can test the assumption of homotheticity by imposing the restriction that $\gamma_{jj} = \gamma_{jk} = \gamma_{kk} \equiv \gamma$ for all $j \neq k$ (Griliches & Ringstad, 1971).

3. Data

This study uses individual-level restricted-access data from the National Education Longitudinal Survey (NELS), first administered nationally to eighth-graders in 1988 by the U.S. Department of Education's National Center for Education Statistics, and updated when the students were tenth-graders in 1990 (U.S. Department of Education, National Center for Education Statistics, 1988–94). The NELS is an integrated set of surveys of students, their parents, their teachers, and their schools. In addition, NELS participants took a battery of tests in mathematics, reading, science, and social studies, developed by the Educational Testing Service, in each survey year. Therefore, it is possible to measure the gains in student academic achievement from the eighth to tenth grades.

I use the approximately 5800 NELS observations with complete records for the student, parent and public school questionnaires in the 1988 survey and the student questionnaire in the follow-up survey (about three-quarters of the possible observations). I dropped observations for a number of reasons. For instance, I dropped observations if the student did not take the tenth-grade follow-up examination, so I was unable to estimate a value-added measure of the student's learning. I dropped those students who changed school districts between the first two waves of the NELS, as I would not be able to identify which school's characteristics contributed to student learning. I also dropped students whose parents did not

⁴ This is not the first paper to estimate translog functions without cost-share equations. Many well-known studies, starting with Corbo & Meller (1979), have estimated translog functions directly, even when cost shares and input prices were available.

⁵ As an alternative, I also constructed bootstrapped standard errors of the elasticities, using the method described by Eakin, McMillen, & Buono (1990). For these bootstraps I drew 2000 samples with 3000 observations apiece. The bootstrapped standard errors closely mirrored the calculated asymptotic standard errors.

complete their questionnaires.⁶ In many respects, the dropped observations closely mirror those used in my sample. However, the dropped observations are disproportionately minority students (28%, rather than 21% in my sample) with lower family incomes (the mean family income of the dropped observations is 8% lower than in the full sample). Furthermore, the dropped observations have, on average, about 4% lower eighth-grade test scores than do the observations used in my sample. Hence, I may be oversampling higher-income, lower-needs students in my analysis. While it is possible that mine is an endogenously selected sample, in which case the point estimates may not be correct, the nature of central finding of this paper is not altered regardless of whether the parameter estimates are affected by selection bias. I know of no way to introduce a selectivity correction into the framework of a translog production function in this application.

I estimated four separate student performance gain equations, with the dependent variables being the change in the student's performance on each of the four tenth-grade NELS examinations—mathematics, reading, science and social studies. My estimation of the four equations was motivated by the possibility that there may be systematic differences in the ways in which school inputs affect learning in different subjects. However, I found that the estimated productivity of school inputs is approximately the same across all four test scores. Therefore, I concentrate only the science results, which are representative of the other three sets of results.⁷

My primary specification (later called the “gains” specification) involves estimation of Eq. (4), in which I include the student's eighth grade test score as a right-hand-side variable. As a check, I also estimate Eq. (4) without the eighth-grade test score as a control variable (although I continue to include the interaction terms between the eighth grade score and all other variables)—the results are qualitatively almost identical to those

presented herein, and are available upon request. As a further sensitivity check, I estimate Eq. (4) using a different data set, *High School and Beyond*, and again find qualitatively identical results to those presented herein.

For computational reasons, as every variable included dramatically increases the number of parameters to be estimated, I must be parsimonious with my variable selection. Therefore, there are certainly variables in the NELS data set that I could have used in my estimation but choose not to. I include four school input variables in my estimation. These variables—the school's student-teacher ratio, the starting salary (in 1988) for a teacher with a bachelor's degree, the number of instructional hours in the academic year (as measured by the number of instructional hours in a school day times the number of school days in the academic year), and the percentage of teachers with master's degrees—are among the most commonly used variables reflecting measured school inputs into the education production process. I recognize that these variables are only weak measures of school quality, and that teacher salaries are a particularly suspect proxy for teacher quality, but these variables are consistent with those generally used in the literature. In addition, numerous researchers, as early as Antos & Rosen (1975) and as recently as Ballou (1996) and Figlio (1997b), have shown that higher-quality teachers also tend to be paid higher salaries, all else equal.

The NELS offers much more detailed measures of school inputs than I include in this study. For instance, one can match students directly to classroom-level inputs (e.g., the size of the student's science class, or the qualifications of that student's teacher). This is a clear advantage of the NELS over prior longitudinal education datasets (see, e.g., Goldhaber & Brewer (1997), who find that unobserved classroom-specific characteristics are strongly related to student test performance, though find that studies such as this one are not biased by excluding such unobservable factors). Therefore, it may seem odd that I am using the more crude measured school inputs in this study. I do so for three principal reasons. The first reason is operational: the translog education production function involves estimating a substantial number of parameters, and sample size becomes an issue. Students in the NELS are matched with one or two, but not all, of their teachers, and not all teacher files are complete records. Going to the classroom-level data, though obviously preferable in one dimension, would reduce my effective sample size by half. Given the number of parameters to estimate, in the tradeoff between more detailed data and sample size, selecting sample size seemed the more compelling choice. My other primary reason for using cruder measures of school inputs is that I want to point out a potential reason for why previous authors tend to find that school inputs do not matter; it seems logical to use the same types of input data as did prior authors. A third reason to use school level data is that

⁶ I recognize that the decision to drop observations with missing parental income values could lead to selection bias: there may be systematic reasons why parents do not reveal their incomes. To get a sense of the severity of this issue, I also imputed the missing income values using parental age, occupation and state of residence, but found that including these observations did not meaningfully change the elasticities that I compute in this paper. I also experimented with including a dummy variable for missing parental income, which also did not seem to make a difference. Thus, rather than rely on imputations or student reports, I choose to omit the observations in which parents do not report their income.

⁷ Out of concerns regarding the potential for measurement error, I also estimated specifications using a composite test score—a weighted average of the four subject area scores. The results are qualitatively similar to those using the science test scores.

while it's true that classroom-level inputs give a better sense of what goes on in one of the student's classes, the student has taken numerous classes in the subject between eighth and tenth grades. Perhaps the school-level input data are not much less representative of the student's actual classroom experiences than are classroom-level data for one of several classes in the subject.

Education production function studies generally include a number of additional variables intended to control for student background and peer characteristics. Although almost all published studies have at least a few of these controls, Summers & Wolfe (1977) have among the most extensive sets of controls. I adopt a close analog to Summers and Wolfe's set of control variables as explanatory variables in my base model. These variables are (obviously unlogged) dummy variables for whether the student is black, or Hispanic, whether the student is male, whether the student has had frequent absences (measured here as at least five in the last month),⁸ whether the student has been offered drugs in school in the last month (as a proxy for a disruptive environment), and whether the student has low-achieving peers. I say that a student has low-achieving peers if the average composite eighth-grade test score in the student's school (excluding the student from the average) is at least half a standard deviation below the mean.⁹ In addition, I control for the percentage of black and Hispanic students in the school. Finally, I include the student's annual family income, up to US\$200 000, as reported by the student's parents.^{10,11}

⁸ The NELS survey does not allow the student to directly enter the number of recent absences, but rather simply identify a category of absence frequency. The five-and-above category seems more appropriate than lumping together all students who have missed school at all.

⁹ This measure of the level of peer quality may be subject to the "reflection problem" described by Manski (1993), in which the peer effects are endogenous. The reflection problem put forward by Manski assumes contemporaneous effects. Therefore, I hope to circumvent the reflection problem by using the average base-year test scores of the student's peers to explain the growth in the student's performance in a later period. This assumption implies that a student's performance in time t will not be a function of the quality of his or her peers in time t , but rather will be related to peer quality at time $t - 1$. According to Manski, if the transmission of peer effects really follows the assumed temporal pattern, the identification problem is alleviated. While this assumption, similar to those made by numerous previous authors (e.g., Borjas, 1992), is not fully convincing, it seems sensible to assume that there would be a lag between peer effects and student performance.

¹⁰ The family's income is truncated at \$200 000; however, less than one percent of the distribution falls in the truncated portion of the distribution. For those with incomes over \$200 000 I assign a family income of \$200 000. Adding a dummy indicator for truncated income values does not appreciably change the results.

I also include two variables not used by Summers and Wolfe that may provide additional information about student achievement.¹² Specifically, I control for the school's eighth-grade enrollment to capture the effects of different school sizes. In addition, I control for the average income (in thousands of dollars) of families with students in the school, as calculated from the NELS parent surveys. This variable is intended to help to deal with the potential endogeneity introduced by the probable event that high-income communities have high-service schools.

I recognize that this set of control variables is not perfect. Variables such as racial composition may be endogenous, and absentee rates may be better characterized as outcomes, rather than inputs. Despite these concerns, I chose this set of explanatory variables primarily to illustrate the effects of different functional forms on education production functions. I specifically *did not* want to use variables not widely used in the existing literature, as I wanted to disentangle the effects of using different functional forms from those of using newer and likely better data. As such, I figured that the set of control variables utilized by Summers and Wolfe, the most widely-cited paper in the literature, would be a reasonable variable list to use. While there is surely a cost associated with the translog in that the number of variables is necessarily limited, when I experiment with adding additional variables (in a linear world) the estimated mean effects of the school resource variables do not seem to change appreciably.

4. Results

In order to gauge whether there exists any evidence of a systematic relationship between school inputs and student achievement, I compute the marginal productivities of each of the four measured school inputs used in this analysis. As mentioned in Section 2, computing these elasticities involves the evaluation of 15 parameters. To conserve on space, and because individual coefficient estimates in the translog are difficult to interpret, I do not report the parameter estimates themselves. A table of the full set of coefficient estimates is available on request.

The first column of Table 1 presents the estimated marginal productivities and asymptotic standard errors of

¹¹ Summers & Wolfe (1977) also control for the student's IQ, a variable unavailable in my data set. However, they mention in a footnote that their regressions do not change much when the student's IQ is excluded from the analysis.

¹² I find that excluding these additional variables and using only my analog to the Summers & Wolfe (1977) control variables results in only small changes in the estimated effects of school inputs on student achievement.

Table 1

Estimated production elasticities of school inputs in translog and Cobb–Douglas specifications (all variables evaluated at their means)

School input	Translog results	Cobb–Douglas results
Student–teacher ratio	−0.139 ^a (0.041)	−0.017 (0.011)
Starting teacher salary	0.134 ^a (0.033)	0.037 (0.024)
Hours in school year	−0.027 (0.019)	−0.046 (0.030)
Percent of teachers with masters degree	0.0004 (0.002)	−0.004 (0.003)

Note: All variables are evaluated at their mean values. These estimated elasticities measure $\partial(\ln y)/\partial(\ln x)$, where y is the tenth-grade science test score and x is the relevant school input. Asymptotic standard errors are in parentheses beneath point estimates.

^aCoefficient estimates are statistically significant at the 5% level.

the four school inputs on the tenth-grade level of student science achievement and the gain in achievement from eighth to tenth grades. In order to construct these elasticities, I evaluate all variables at their means. Later in this Section I present estimates of the elasticities of production at different points in the distributions of selected variables.

There is an apparently small but systematic relationship between certain measured school inputs and student performance. The estimated production elasticity of student–teacher ratios, evaluating all variables at their means, equals −0.139, and is statistically distinct from zero. This result suggests that, all else equal, students attending schools with lower student–teacher ratios fare somewhat better than those attending schools with more students per teacher. Specifically, this result suggests that a 10% increase in the student–teacher ratio is associated with about 1.4% lower tenth-grade achievement, holding constant the level of achievement in eighth grade.

I also observe precise estimates of the marginal effects of starting teacher salaries on student achievement. The estimated elasticity of teacher salaries is 0.134 in the gains equation. Therefore, the results suggest that students attending schools with higher-paid teachers perform at a higher level of competence than students taught by lower-paid instructors, *ceteris paribus*. In particular, the point estimates indicate that increasing starting salaries by 10% will result in 1.3 higher tenth grade achievement, holding eighth grade performance constant.

The estimated effects of the percentage of teachers with master’s degrees are extremely small (less than 0.01 in absolute value) and are not precisely estimated. This result suggests that the very high premia paid for master’s degrees in most school system salary schedules may not be buying much in the way of increased student success.

Unlike student–teacher ratios and teacher salaries, for which the marginal productivities in the gains and levels specifications were of the same sign, the estimated production elasticities of instructional hours are negative, but statistically indistinguishable from zero, in the gains specification. (This is the one place where the gains

results differ substantively from a levels-type specification, in which the relationship between instructional hours and achievement levels is positive and significant, suggesting that much of this benefit must accrue in the grades prior to eighth grade, if at all.)

4.1. Effects of school inputs at different levels of other variables

The production elasticities of school inputs are different for every observation. In order to provide an idea of how these elasticities vary over different levels of selected variables, I present in Figs. 1 and 2 graphs of the estimated marginal productivities of student–teacher ratios and starting teacher salaries over the distribution of six variables—the student–teacher ratio, starting teacher salary, instructional hours, school eighth grade enrollment, average family income in the school, and the student’s initial (eighth grade) achievement level. As before, I evaluate all variables at their means except for the one that I allow to vary. Because the estimated effects of the percentage of teachers with master’s degrees are extremely small and not precisely estimated, I focus the rest of my discussion solely on the estimated effects of the other three school inputs.

The numbers on the vertical axis of these figures represent the marginal effects of the relevant school input, while the numbers on the horizontal axis represent the cumulative distribution of the variable that I allow to change. Therefore, the lines on the graph represent how the estimated effects of school inputs change as we go from low levels of the variable of interest (at the left side of the graph) to high levels of the variable (toward the right).

Fig. 1 illustrates these marginal effects of changing the student–teacher ratio in the gains specification. The most pronounced differences in the marginal effects of student–teacher ratios occur as the eighth grade school enrollment changes. In small schools, the marginal effects of student–teacher ratios can be less than half the size (in absolute terms) of the marginal effects seen in large schools. The marginal effects of student–teacher

Fig. 1. Effects of student–teacher ratios across different levels of variables.

ratios also increase modestly in magnitude as instructional hours and school income levels increase. These interaction effects are statistically significant at the 5% level. The effects of student–teacher ratios seem to be largely independent of the level of the student–teacher ratio, starting salary level, or initial achievement level.

Fig. 2 repeats the same exercise as Fig. 1, but presents the estimated effects of starting salaries at different levels of the six variables. Here, the most pronounced changes occur when the school's income level or starting salaries vary. The results suggest that, evaluating all variables at their means, the higher the starting salary, the greater the marginal effect of further increasing starting salaries. On the other hand, the higher the average family income in the school, the lower the estimated marginal effects of teacher salaries. There is also a modest positive relationship between the student–teacher ratio and the marginal effect of starting salaries, and a negative relationship between initial achievement and the marginal effects of starting salaries. All of these interactions are significant at at least the 10% level. These results suggest that starting salary increases would be most effective in schools with low family incomes, large classes, and low initial achievement levels.

Fig. 2. Effects of starting salaries across different levels of variables.

4.2. How substitutable are school inputs?

I have found limited evidence suggesting that, say, lower student–teacher ratios and higher teacher salaries are associated with improved student performance. From a policy perspective, it is useful to gauge the substitutability of school inputs in the education production process. That is, if the student–teacher ratio rises, how much of an increase in teacher salaries, say, would be required to maintain the same level of student achievement, all else equal? When I evaluate all variables at their means, I find that a 10% increase in the student–teacher ratio would require a 10.4% increase in teacher salaries to maintain the same level of student achievement, *ceteris paribus*.

But is this tradeoff the same for all students, or all schools? One previously-stated benefit of the translog functional form is that the elasticities of production and substitution differ for every observation. I therefore compare these marginal rates of transformation across different levels of variables. First, compare “high service” schools (defined here as schools with fifth-percentile student–teacher ratios (about 11) and 95th-percentile starting teacher salaries (about \$22 300)) to “low service” schools (schools with 95th-percentile student–teacher ratios (about 25) and fifth-percentile teacher salaries (about \$14 600)). For high service schools, a 10% increase in student–teacher ratios requires a 9.65% increase in salaries, while in low service schools the 10% increase in student–teacher ratios would require 11.1% higher salaries. Therefore, the results suggest that improving achievement through salary increases works better for high-service schools, while changing the student–teacher ratio would work better for low-service schools.

These tradeoffs also vary considerably with school enrollment. For low-enrollment schools (fifth-percentile, with a per-grade enrollment of 46 students), a 10% increase in the student–teacher ratio requires 7.4% higher salaries to attain the same level of student achievement, while for high-enrollment schools (95th-percentile, with a per-grade enrollment of 515 students), the same 10% increase in the student–teacher ratio would be offset by a 12.3% increase in teacher salaries. Hence, salary increases seem more advantageous in small schools, while student–teacher ratio decreases appear more advantageous in large schools.

Schools with low average family incomes (fifth-percentile, with average family income of about \$18 100) require 9.4% higher salaries to compensate for a 10% increase in the student–teacher ratio. On the other hand, schools with high average family incomes (95th-percentile, with average family income of about \$71 600) require 11.4% higher salaries to offset a 10% increase in the student–teacher ratio. Therefore, the relative advantage of increasing school quality through lower

student–teacher ratios increases with the incomes of families whose children attend the school. While simply illustrations, these results suggest that the tradeoffs between school inputs can vary considerably at different levels of school and family inputs. The results indicate that it should be possible to identify “optimal” mixes of school inputs to maximize student achievement at least cost. While I do not attempt to carry this out in this paper, I intend to explore this issue in future research.

4.3. *Does money matter?*

These results provide new evidence of a systematic relationship between measured school inputs and student performance. However, there is a question as to how substantial these effects are. Although the results are statistically significant, they seem small in magnitude. Ten percent increases in student–teacher ratios apparently result in only 1% lower changes in student performance, or less than 2% decreases in cumulative student achievement. Ten percent increases in starting salaries are also associated with small changes in student performance—about 1.4% higher performance growth from the eighth grade level, or less than 3% increases in cumulative student achievement. Only instructional hours seem to have very substantial effects on student achievement, with 5% increases in cumulative student achievement associated with 10% increases in instructional hours. However, the estimated effects of instructional hours on gains in achievement are negative (though not statistically significant). And 10% changes in these variables are very large—the U.S. Department of Education’s Common Core of Data show that 10% decreases in student–teacher ratios, for instance, have occurred very rarely since these data have been collected. In general then, it seems that even very large changes in school inputs, well beyond what any single school district would consider, do not have particularly large effects on student achievement. Therefore, although I conclude that school inputs are likely related to student performance, I also argue that the school inputs that I have measured may not matter much.

In addition, there is reason to believe that my estimated elasticities (and those of prior authors) may still be overstated. Higher schooling inputs are correlated with higher levels of observable parental inputs—the parents of students attending high-input schools tend to be better educated and make higher incomes than the parents of students attending low-input schools. If there are unobservable student skills and talents that are correlated with the observable features of parents, then this would likely bias upward education production function estimates of the effects of measured school inputs on student achievement. Therefore, my already modest estimated effects of measured school inputs on student achievement, and those of all prior authors, may actually still be too high.

5. Data or method?

I have observed that, at least for the NELS data, the functional form assumptions of the existing education production literature may be overly restrictive. When I estimate a translog education production function, I find evidence of a systematic relationship between school inputs and student performance.

Because I estimate education production functions with both a new data set and a new estimation method, it is not clear whether my results are driven by the estimation method or by the choice of data. In order to determine whether the results presented above would be found in the NELS data regardless of the functional form specification, I estimate a traditional logarithmic Cobb–Douglas model using the same explanatory variables as in the translog specification described in the previous section. This specification can be written in the notation from Eq. (4) as $\ln y_t = \alpha_0 + \sum_k \beta_k (\ln X_{kit} - 1) + \alpha_1 \ln y_{t-1}$.

I report the estimated production elasticities of school inputs resulting from the Cobb–Douglas estimation in the second column of Table 1. As with the translog estimates, I evaluate the Cobb–Douglas results in two stages. First, I consider the precision of the estimated effects of school inputs on student achievement. The results of the Cobb–Douglas specification differ substantially from those of the translog model. Although the Cobb–Douglas elasticities have the same estimated signs as the translog elasticities, never do I observe a point estimate that is statistically distinguishable from zero.¹³ Therefore, I can conclude that analysis of the NELS data within the framework of a traditional Cobb–Douglas education production function would provide additional support to the widely-stated argument that there exists no systematic relationship between school inputs and student achievement.

The second pass at analyzing the translog results involved examining the magnitudes of the estimated elasticities. While the translog elasticities seem rather small, they are very large in comparison to the Cobb–Douglas elasticities. For instance, the estimated Cobb–Douglas elasticity of student–teacher ratios on student achievement is -0.017 , in contrast to -0.139 in the translog specification. The estimated Cobb–Douglas productivity elasticity of starting salaries is 0.037 , but 0.134 in the translog specification. Indeed, only in the case of the percentage of teachers with master’s degrees are the Cobb–Douglas and translog elasticities consistently of comparable magnitude, but as mentioned before, this magnitude is trivial. Therefore, I can conclude that at

¹³ The results using the other test scores are generally the same, except that I observe a statistically significant estimate of the effect of student–teacher ratios on social studies achievement.

Table 2
F-test statistics for various hypotheses: science scores

Hypothesis	Gain in science achievement	Tenth grade science score	1% critical values ^a (gain/10th grade)
(1) Cobb–Douglas production	35.255	6.730	2.14/2.15
(2) Group-wise additivity	20.857	3.775	1.84/1.87
(3) Homotheticity	32.731	6.110	2.09/2.10

^aThe 1% critical value is taken from the F -distribution with J degrees of freedom in the numerator and $(n - K)$ degrees of freedom in the denominator. J is the number of parameters in the restricted (null hypothesis) model, K is the number of parameters in the full model, and n is the number of observations. The null hypotheses tested above are $\gamma_{jj} = \gamma_{jk} = \gamma_{kk} = \theta_j = 0$ for all $j \neq k$ for the Cobb–Douglas test, $\gamma_{jk} = 0$ for all $j \neq k$ for the additivity test, and $\gamma_{jj} = \gamma_{jk} = \gamma_{kk} = \gamma$ for all $j \neq k$ for the homotheticity test.

least for the NELS data, estimating the Cobb–Douglas specification results in considerably smaller estimated effects of school inputs than are found when the education production function is modeled as a translog function.

5.1. Hypothesis testing

When I estimate a translog education production function, I observe a statistically significant relationship between school inputs and student achievement. Is the complicated translog functional form necessary, or are traditional estimation methods sufficient? I test the maintained functional form assumptions of the existing education production function literature. If the translog function does not provide significant informational advantages over the less complicated estimation strategies, the utility of estimating a flexible functional form is diminished.

I report the results of various hypothesis tests in Table 2. First, consider the hypothesis that the true education production function is Cobb–Douglas, the common assumption used in the existing literature. To do this, I conduct an F -test in which I compare the explanatory power of a restricted model (here, the Cobb–Douglas) to that of the full translog model. The first row of the table presents the test statistics for the levels and gains specifications, as well as the corresponding 1% critical values. I observe that the hypothesis that the production function is Cobb–Douglas can be rejected easily at the 1% or better level.¹⁴ This result suggests that, at least for the NELS data, the commonly-used Cobb–Douglas assumption is too restrictive.

Next, I test the other standard assumptions of group-wise additivity and homotheticity by making the restrictions described in Section 2. The test statistics and critical values for these tests are reported in rows two and three of Table 2. As with the test for Cobb–Douglas pro-

duction, I find that I can reject the null hypotheses that the education production function is either additive or homothetic at the 1% or better level. Hence, these results suggest that authors in the existing individual-level education production function literature may not have been employing a sufficiently flexible functional form in their estimation. At the very least, the results of the hypothesis testing suggest that an additive or homothetic functional form is not appropriate for the NELS data used in this study, and that the translog functional form provides a more accurate representation of the education production relationships.

I find that the Cobb–Douglas is never a particularly good approximation of the education production function. But there are places where the Cobb–Douglas performs better than others. For instance, these test statistics are lowest in wealthy school districts (those with average family incomes of more than one standard deviation above the mean) and in small schools (those with per-grade enrollments of 100 students or fewer). The statistics are highest in relatively poor school districts (those with below-average family incomes), large schools, and schools with large minority populations. It should be noted, however, that no matter how the data are sliced, I can always reject the Cobb–Douglas specification in favor of the translog at at least the 1% level.

6. Conclusion

The lack of statistically significant relationships between measured school inputs and student academic achievement found in the great majority of published research in the education production function literature has led many researchers, most notably Hanushek (1986, 1991), to argue that there is no discernable relationship between school inputs and student success. This study finds evidence that, at least for the NELS data (and also, for a comparable specification using HSB), these results may be driven in part by the functional form assumptions of the existing literature. Specifically, I find that the standard assumptions of additivity and homotheticity

¹⁴ The results for this and all other hypothesis tests reported hold for all four test scores.

employed by the existing individual-level literature may be overly restrictive.

When I estimate an education production function with the flexible translog functional form, I find evidence of a systematic relationship between measured school inputs and student achievement. Specifically, I find that higher student–teacher ratios are associated with lower student performance and that higher starting teacher salaries are positively related to increased student performance, especially in schools with low family incomes, large classes, and low initial achievement levels. More important than the statistical significance of the results, however, are the magnitudes of the results. I compare my translog findings with traditional Cobb–Douglas education production function using the same data and set of variables and find that the translog production function yields considerably larger estimated effects of school inputs than does the Cobb–Douglas function. This result suggests that traditional education production functions may underestimate the effects of school inputs on student achievement.

I cannot make the argument that money matters simply because I find precisely estimated evidence of a relationship between school inputs and student achievement. In order to determine whether variation in measured school inputs makes much of a difference, one must also evaluate the magnitudes of the estimated effects. My results suggest that simply lowering pupil–teacher ratios, increasing starting salaries of teachers, or lengthening the school day (or year) will not dramatically improve student achievement. Although the evidence implies that student achievement will increase measurably by such an expenditure, the benefits of small improvements in student performance resulting from even very large increases in measured school inputs are likely offset by the enormous costs of bringing about such a change. Instead, the results suggest that students may benefit more by having schools follow the advice of Hanushek (1994) and change the ways in which education is delivered in schools.

There surely exists considerable variation in school quality that is not captured by differences in the measured school inputs used in this study. For instance, the percentage of teachers with master's degrees and the starting teacher salaries are only crude indicators of teacher quality. Future research should attempt to use as detailed information as possible about specific characteristics of the school, its programs, and its teachers, so that the researcher can better determine the particular mechanisms through which schools influence student achievement. In addition, these results suggest that researchers estimating education production functions should focus more attention on the specific functional form assumptions that they employ.

Acknowledgements

I am grateful to seminar participants and other colleagues at the University of Oregon, University of Wisconsin, and Virginia Tech, as well as two anonymous referees. I owe particular thanks to Rick Hanushek, Bob Haveman, Andy Reschovsky, Karl Scholz, Joe Stone, and Bobbi Wolfe for their helpful advice. I acknowledge research support from the U.S. Department of Education, the University of Oregon Foundation, and the Wisconsin Alumni Research Foundation. All errors are my own.

References

- Antos, J., & Rosen, S. (1975). Discrimination in the market for public school teachers. *Journal of Econometrics*, 3, 123–150.
- Ballou, D. (1996). Do public schools hire the best applicants? *Quarterly Journal of Economics*, 111, 97–134.
- Borjas, G. (1992). Ethnic capital and intergenerational mobility. *Quarterly Journal of Economics*, 107, 123–150.
- Callan, S., & Santerre, R. (1990). The production characteristics of local public education: a multi-product and input approach. *Southern Economic Journal*, 57, 468–480.
- Christensen, L., Jorgenson, D., & Lau, L. (1971). Conjugate duality and the transcendental logarithmic production function. *Econometrica*, 39, 255–256.
- Corbo, V., & Meller, P. (1979). The translog production function: some evidence from establishment data. *Journal of Econometrics*, 10, 193–199.
- Downes, T., & Figlio, D. (1998). *School finance reforms, tax limits, and student performance: do reforms level-up or dumb down?* Discussion paper. University of Wisconsin-Madison: Institute for Research on Poverty.
- Duncombe, W. (1992). Costs and factor substitution in the provision of local fire services. *Review of Economics and Statistics*, 74, 180–184.
- Eakin, B. K., McMillen, D., & Buono, M. (1990). Constructing confidence intervals using the bootstrap: an application to a multi-product cost function. *Review of Economics and Statistics*, 72, 339–344.
- Ehrenberg, R., & Brewer, D. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review*, 13, 1–17.
- Eide, E., & Showalter, M. (1998). The effect of school quality on student performance: a quantile regression approach. *Economics Letters*, 58, 345–350.
- Figlio, D. (1997). Did the “tax revolt” reduce school performance? *Journal of Public Economics*, 65, 245–269.
- Figlio, D. (1997). Teacher salaries and teacher quality. *Economics Letters*, 55, 267–271.
- Goldhaber, D., & Brewer, D. (1997). Why don't schools and teachers seem to matter? *Journal of Human Resources*, 32, 505–523.
- Griliches, Z., & Ringstad, V. (1971). *Economics of scale and the form of the production function*. Amsterdam: North-Holland.
- Gyapong, A., & Gyimah-Brempong, K. (1988). Factor substitution, price elasticity of factor demand and returns to scale

- in police protection: evidence from Michigan. *Southern Economic Journal*, 54, 863–878.
- Gyimah-Brempong, K., & Gyapong, A. (1992). Elasticities of factor substitution in the production of education. *Economics of Education Review*, 11, 205–217.
- Hanushek, E. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141–1177.
- Hanushek, E. (1991). When school finance “reform” may not be good policy. *Harvard Journal of Legislation*, 28, 423–456.
- Hanushek, E., & Members of the Panel on the Economics of Educational Reform (1994). *Making schools work*. Washington, DC: Brookings Institution.
- Hanushek, E., Rivkin, S., & Taylor, L. (1996). Aggregation and the estimated effects of school resources. *Review of Economics and Statistics*, 78, 611–627.
- Heady, E., & Dillon, J. (1961). *Agricultural production functions*. Ames, IA: Iowa State University Press.
- Jimenez, E. (1986). The structure of education cost: multiproduct cost functions for primary and secondary schools in Latin America. *Economics of Education Review*, 5, 25–39.
- Kim, H. Y. (1992). The translog production function and variable returns to scale. *Review of Economics and Statistics*, 74, 546–552.
- Manski, C. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, 60, 531–542.
- Summers, A., & Wolfe, B. (1977). Do schools make a difference? *American Economic Review*, 67, 639–652.
- U.S. Department of Education, National Center for Education Statistics (1988–94). *National education longitudinal survey of 1988*. Washington, DC: U.S. Department of Education.