

# Education Production Functions Using Instructional Time as an Input

Dennis Coates  
Department of Economics  
UMBC  
1000 Hilltop Circle  
Baltimore, MD 21250

coates@umbc.edu

Draft date: February 25, 1998

Paper prepared for the meetings of the Public Choice Society held in New Orleans, Louisiana, March 13-15, 1998. I wish to thank to Rense Lange of the Illinois State Board of Education for helpful explanations regarding the Illinois Goals Assessment Program test scores.

## Abstract

This paper contributes to the literature on the effectiveness of schools. The analysis is unique in five respects. First, the data employed include the minutes of instruction per five day week in each of four subjects for all the public schools in Illinois. Few education production function papers have any information on the amount of instruction students receive in a given subject. Second, the theory section of the paper argues that class size should interact with student, teacher, and school characteristics as well as with the instructional times. The empirical analysis tests this hypothesis. Third, the interactions show that the small marginal effects of class size or of teacher qualifications may result because the harm of larger class sizes is undone by better trained personnel, and the benefits of better trained teachers is undone by large class sizes. Fourth, the data span three years enabling the use of panel data techniques. Individual school specific effects are estimated and found to be strongly significant. Moreover, in the presence of these effects, class size always has the theoretically predicted sign and is highly significant. Fifth, the time series aspect of the data allows for tests of the comparability of test scores over time. If the test scores are to be compared over time as a guide to school performance and improvement, then those scores should not differ significantly over time. The results reveal that year to year differences can be substantial and statistically significant.

Education reform is a hot topic among politicians, academics and even around the dinner table these days. The issues are complex and important. One focus of the debate is on the level of resources appropriated by state and local politicians to the public school system. This debate has come to be identified with the question “Does money matter?”

There seems to be no consensus on this issue. On the one hand, researchers (Hanushek, 1996, for example) point out that school expenditures per pupil have risen many times over in the last thirty or so years, but educational achievement, as measured by scores on standardized tests, has not. Others point out that much of this expenditure growth has been directed at special education students, not the average student. When this is accounted for, spending per pupil has risen, though not nearly so rapidly.

Still others (Hedges and Greenwald, 1996) chime in by suggesting that over this same period of time family structures have deteriorated, more children are in single-parent households, more are living in poverty, more have less connection to books or well-educated people, than ever before. Consequently, the task of educating children has grown more difficult over time. More resources are needed just to stay in the same place given the deterioration in other circumstances.

Much of the research that underlies these conclusions is rooted in a production function paradigm, familiar to economists but not nearly so well-known among non-economists. Given the work that economists have done in this area, however, just how familiar economists are with the production model may also be questioned. This paper goes back to the basics of production theory to motivate a new empirical approach to the estimation of an education production function. Specifically, the analysis here recognizes that production functions are flow, not stock, relationships. Consequently, any attempt to model the process of education must account not

only for static characteristics of the schools and teachers, the children and their families, but also for the continual application of resources to production.

The results of rethinking the production relationship indicate that the functional forms and variables included in the typical education production function equation are lacking. The results indicate that the amount of time spent in instruction is important in determining average test scores. Moreover, the effect of instruction time varies with class size; a given instructional time has smaller impact on test scores when classes are big than when classes are small. In addition, increased class size has negative consequences for average test scores once the instructional times variables are included.

Pooling the data reveals that year and school specific effects are very important. That year effects are important is troublesome for any program which professes to use aggregate test scores as measures of performance because significance of the year effects indicates that whether a school appears to have improved or not is in part determined by fluctuations in the difficulty of the test from year to year. This also raises a question about the use of the “value added” approach to education production functions because what appears to be value added may have more to do with the tests and their grading than with what students have learned. Moreover, using the pooled data with year and school specific effects reveals a very strong effect of average class size with the expected sign.

The next section of the paper returns to the production function as typically defined and described by microeconomists. This analysis suggests some respecifications of the empirical model which are implemented on data from the state of Illinois. The data are described in section XX and the results in YY. A conclusion follows.

## I) The Production Function

Careful discussions of production functions describe the function as relating the amount of output produced per unit of time to the amount of resources utilized per unit of time in producing the output. Having machinery in the factory sit idle produces nothing whereas having that machinery operating results in generation of output. Idle machinery supplies no services per unit of time, operating machinery does. Similarly, workers produce nothing when not actually doing their jobs.

At the same time, the services per unit of time from one type or vintage of equipment or machinery might be very different from those of another type or vintage. Workers with different skills or experience might also provide different levels of labor services in a given amount of time. It is also likely that the ability to get services from some machinery or to use some materials effectively, depends upon the skills and experience of the worker. For example, the quality of the framing for a building that a new worker, one without prior construction experience, produces with warped and knotty wood may be quite different from the quality produced by a master carpenter with the same wood. And the time it takes the former to produce the same structure may be substantially more than what it takes the latter.

These examples suggest that when discussing production, three concerns regarding resource usage are likely to be quite important. The first of these is the time spent in the production process. The second concern is with the quality of the resources used in the process. Finally, the effects of the quality of resources and time spent by the resources are interdependent.

Few studies of education production functions address these issues. Exceptions to this

include Fredrick and Walberg (1980)<sup>1</sup>, Walberg, Fraser, and Welch (1986) and Betts (1997). One possible reason for this omission in most studies is that data on the time spent in instruction is rarely available in precise form. For example, in Walberg, Fraser, and Welch (1986) instructional time is measured as the number of semesters of science a student took in grades 9 through 12. Betts (1997) uses the teacher's report of the number of hours of math class as his measure of instructional time. This paper has the advantage of having instructional times by subject matter, english, mathematics, science and social studies, measured as the average minutes per five-day week in the third grade classrooms for each of the Illinois public schools.

Several studies have, however, suggested that a nonlinear relationship exists between inputs into the education process and the output of that process. Summers and Wolfe (1977), for example, argue that the impact of teacher quality and class size differ by the race of the student. Ferguson (1992) suggests that the effect of class size is nonlinear, that there are discrete jumps in the effect of class size not constant marginal impacts. Ferguson and Ladd (1996) take account of these types of nonlinearities as well.

Betts (1996) examines the effects of longer homework assignments on standardized test scores in mathematics. He notes that homework "appears to be a much more significant predictor of gains in math test scores than the more standard measures of school inputs such as class size and teachers' credentials." Moreover, the analysis indicates that the marginal productivity of homework in producing test scores is constant, at least within the range of homework assigned

---

<sup>1</sup>Fredrick and Walberg (1980) discuss review a broad array of studies of the relationship of time to learning. For example, attendance rates, number of absences and number of times tardy are reflective of time inputs in the education process. More relevant for the issue raised here, their survey also examines the effects of hours of classes, minutes of study, and minutes or proportion of time on task.

observed in the data. The clear inference from Betts' analysis is, therefore, taken to its logical conclusion, that an inexperienced poorly trained "teacher" facing a class of several hundred need only assign several hours of homework a night to produce students with high test scores.

Interestingly, these studies do not address the possible interaction of class size with teacher characteristics.<sup>2</sup> Yet intuition suggests that better trained or more experienced teachers may be better able to educate in large classes than are poorly trained or inexperienced teachers. Ehrenberg and Brewer (1994) find different effects of teacher characteristics by race and, indirectly, by class size and other school and student variables besides. They note that in the High School and Beyond data, a student that was surveyed as a sophomore may have dropped out before being surveyed as a senior. Not accounting for these possibly non-random withdrawals from the sample could bias the results of the education production function estimates. Consequently, they correct for sample selection bias using the inverse Mills' ratio constructed from a probit equation predicting continuation in the sample from sophomore to senior surveys. Of course, this correction introduces the teacher, school and student variables into the education production function in a highly non-linear fashion. Ehrenberg and Brewer (1994) find the inverse Mills' ratio to be statistically significant, suggesting sample selection biases exist. At the same time, they note that the magnitudes and statistical significance of the other variables are not much affected. An alternative interpretation is that the significance of the inverse Mills' ratio indicates

---

<sup>2</sup>Hanushek, Rivkin, and Taylor (1996) explicitly ignore interactions because they find that the effects of race or student scores on earlier tests do not influence the effects of school characteristics. Krueger (1997) uses experimentally generated data on school children from Tennessee to assess the importance of class size. He makes no control for interactions of teacher characteristics and class size. This may be especially important in his analysis because while small classes are shown to be beneficial, regular size classes with full-time teachers' aides show no differential effects over regular size classes without aides. Krueger hypothesizes that this may be due to measurement error because even the regular classrooms routinely have part-time aides. Interaction of teacher characteristics with classroom size might enable one to assess this possibility.

substantial nonlinearity and interaction among the explanatory variables. To assess the marginal impact of any variable is, therefore, more involved than simply performing a t-test on its coefficient.

Ludwig and Bassi (1997) test for specification errors using the Ramsey RESET test in their analysis of NELS data on the relationship between school spending and student achievement. They find evidence that the typical “gain equation” is misspecified, though they cannot conclude whether the problem is omitted variables, incorrect functional form or endogeneity. The RESET test, recall, uses the fitted values of the dependent variable in higher order terms as regressors. If these terms are statistically significant, that indicates that the error from the original equation contained the influence of the explanatory variables in highly nonlinear fashion. From the perspective of this paper, such a finding supports the idea that the typical equation does not account sufficiently for the interactions of the inputs in the production process.

Ludwig and Bassi use a feasible generalized least squares estimator in which the regression errors from a first stage are used to construct the first-order correlation in the error terms. This correlation is then used as a weight in the differences of differences in both the dependent and independent variables. Since the errors are constructed from the dependent variable and the independent variables in a highly nonlinear fashion, one might interpret the analysis as favoring a highly nonlinear education production function, one in which current and past values of the input variables interact in complicated ways. Ludwig and Bassi find no evidence, in this specification, of misspecification or omitted variables.

The discussion above suggests that an education production function ought to include instructional time, characteristics of the teachers, students, school and community and interaction



terms between time and the other variables. The interaction terms expressly address the issue of producing education with “inputs” of variable quality. More time spent in instruction may be an approach to counteract the effects of large class size, say. Better trained or more experienced teachers may be able to impart the knowledge to children in less time than less trained or inexperienced teachers.

Additionally, instructional time in one subject may have beneficial effects in other subjects. Or it might be detrimental. In other words, this analysis may be able to provide guidance to the type of internal restructuring that many see as the best hope for improving the state of public education in the United States.

Some work has been done on the effects of different production techniques on learning. For example, Julian Betts has working papers on the role of homework in improving school quality, the effects of ability tracking, and on grading standards.<sup>3</sup> In the paper on the role of homework, Betts finds that additional homework and additional time in class each have significant effects on test scores. More interesting still, he finds that one cannot reject equality of the effects. That is, more time spent on math homework improves scores on standardized tests by the same amount as more time in class.

## II) The Model

This section describes the research strategy and the empirical specification of the educational production function. The estimation strategy is to begin with models most like those typically found in the literature, then to progress to more complicated models which more

---

<sup>3</sup>See Betts’ papers for citations to similar research.

accurately reflect the true nature of production.

The first model estimated is given by the following equation:

$$S_{jtk} = \alpha + X_{jt}^d \gamma + X_{jt}^s \mu + \epsilon_{jtk} \quad (1)$$

Where  $S_{jtk}$  is the average score on the Illinois Goal Assessment Program test of subject matter k in school j in year t,  $X_{jt}^d$  is district specific variables for school j in year t,  $X_{jt}^s$  is school specific variables for school j in year t,  $\gamma$  and  $\mu$  are vectors of parameters to be estimated,  $\alpha$  is the intercept, and  $\epsilon$  is an iid disturbance term. I will describe the data in detail below. To help fix ideas, note that I will estimate equation (1) for math, reading and writing, for each of the three years 1994-95, 1995-96, and 1996-97.

The second model estimated pools the data across years. This makes two possibilities. The first possibility is to restrict all slope coefficients and the intercept to equality across years. This approach implies that the relationship between the X variables and test scores is invariant over time and that the test scores are, on average, the same across years. Second, one can allow the intercept to vary by year but force the slope coefficients to be constant across years.

This second model addresses an interesting question which has been the subject of recent research. Ferguson and Ladd (1996) and Ladd, Roselius, and Walsh (1997) have examined the methods states use to evaluate progress within schools. These models are built on the assumption that the test scores are comparable over time; that is, a given score in one year means the same as that same score in a subsequent or previous year. If exams are successfully written so that scores are comparable over time, then the intercept should be constant over time.

The panel nature of the data makes one further extension of the basic model possible.

Since each school may appear in the data from 1 to 3 times, school specific intercepts are estimable. The final specification of the typical model then will include both year and school specific intercepts. The school specific intercepts may be thought of as controls for specific principals, to the extent the principal does not change over time, or an infinite variety of other possible characteristics of the school's environment.

The traditional model is then extended to include measures of the time spent on instruction in each of four different subjects and interactions of the instructional time variables with the average class size. The argument for including these variables is simply that student learning of any subject should be related to the time spent on teaching it. In the extreme, one would not expect students who have never been exposed to a foreign language to fare well on a test on the grammar of that language. Similarly, schools which devote many hours each week to instruction in mathematics should find the students performing better on math exams than schools whose students get little math instruction.

This model may be written as:

$$S_{jtk} = \alpha + X_{jt}^d \gamma + X_{jt}^s \mu + MPD_{jt} \beta + (MPD_{jt} * X_{jt}^s) \delta + \epsilon_{jtk} \quad (2)$$

where  $MPD_{jt}$  is a vector of instructional time variables and  $X_{jt} = (X_{jt}^d, X_{jt}^s)$ . The theory suggests that the element of the parameter vector  $\beta$  most closely related to the subject matter of test k have a positive sign. Theory does not have unambiguous implications for the signs of the other elements of  $\beta$ . For example, what role increased instructional time in social studies will have on math test scores is not clear. On the one hand, more social studies instruction means less time for math, all other things equal, and lower math test scores. On the other hand, more time in social

studies instruction might help students maintain focus on the subjects by peaking their interest or breaking up the monotony.

The vector of parameters  $\delta$  captures the interaction effects. For example, one element in the vector measures the indirect effect of added science instructional time at different average class sizes. Consequently, to determine the effects of average class size requires knowing both the coefficient on class size but also each of the coefficients on the instructional time-class size interaction variables and the instructional times. The partial effect of class size on test scores would be:

$$\frac{\partial S_{jtk}}{\partial \text{Average Class Size}_{jt}} = \mu_{ACS} + MPD_{jt}^{ACS} \delta_{ACS} \quad (3)$$

where the ACS sub and superscripts indicate that only those elements of the vectors which involved the average class size variable are included.<sup>4</sup> The effect of instructional time on test scores is:

$$\frac{\partial S_{jtk}}{\partial MPD_{jt}} = \beta_{MPD} + ACS_{jt} \delta_{MPD} \quad (4)$$

where MPD is any of the instructional time variables, without loss of generality.

The individual elements of the vector  $\delta_{ACS}$  are the cross partial derivatives of the instructional time variables on the marginal effect on test scores of increased average class size or,

---

<sup>4</sup>If average class size enters in quadratic form then equation (3) would also include a term which is the coefficient on ACS squared multiplied by 2 times the coefficient on the variable.

by Young's theorem, the effects of increased class size on the marginal effect (productivity), measured as gains in test scores, of additional instructional time  $\delta_{MPD}$ . The individual  $\delta_{MPD}$  are, of course, simply the elements of  $\delta_{ACS}$ . Intuition suggests that the  $\delta_{MPD}$  should be negative; as average class size rises, the marginal benefit (productivity) of instructional time declines. Said another way, given two classes of different size but the same instructional time, one would expect an increase in instructional time to have a greater beneficial effect in the smaller class. Simply put, each student in the smaller class gets a larger share of the added time than each student in the larger class, and so derives greater educational benefit from that time than the students in the large class.

Indeed, intuition suggests that class size ought to interact with most of the other variables. More qualified teachers, as measured by education or experience, might be expected to do better with large classes than less well trained or inexperienced teachers. Large classes might not have the same impact in schools with low attendance as in schools with high attendance; or the effect of class size may differ between low and high income schools, between schools with predominantly white student bodies and those with predominantly nonwhite students, and so on. In the most extensive model, this analysis includes interactions between average class size and every other variable except the school and the year specific intercepts.

### III) The Data

The data used in this analysis comes from the Illinois Goal Assessment Program in the years 1994-95, 1995-96, and 1996-97. In 1985 Illinois passed education reform laws requiring

the improvement of education through setting of “clear learning outcomes, assessing student attainment of the learning outcomes, and developing plans for school improvement”. The assessment portion of this reform, the IGAP program, started in 1988. The IGAP program was designed with four purposes in mind. These are:

1. Help the state evaluate the extent to which public schools in Illinois are meeting the state goals for learning;
2. Describe how schools and districts perform in comparison to the state and nation;
3. Chart progress of schools, districts, and the state over time; and
4. Generate information which can be used for school accountability, policy making, and school improvement.

The program tests 3rd, 6th, 8th, and 10th grades each year in reading, mathematics, and writing, and 4th, 7th and 11th grades in science and social studies. The results of the test are reported to the school and the district in summary form and individual information is provided to the school or district and the parents of each child. Classroom summaries are not prepared.

The Illinois State Board of Education evaluates each school by the share of its IGAP scores that fall into three performance levels. The levels are “does not meet”, “meets”, or “exceeds” the state goals. However, “meets state goals” is further divided into high, middle and low levels. It is important to note that the evaluations apply to a “school building” and depend, therefore, on all the IGAP scores in the building. So, for example, the evaluation of a school which houses first through seventh grades would be based on the reading, writing, and math scores of the third, and sixth graders and the science and social studies scores of the fourth and

seventh graders. The IGAP scores are described in more detail below.

The schools in the sample for this analysis come from over 800 districts in Illinois, totaling more than 2500 schools in which third grade is taught at least one year between 1994-1995 and 1996-1997. The full sample comprises 6807 observations over the three year period. About 1300 of these observations come from schools in the City of Chicago School District. No other single district contributes more than 125 observations.

Variable definitions are provided in Table 1. Descriptive statistics are in Table 2 for the full sample, all schools in all three years combined, and for the schools in each of the three years. The general rule of thumb is that any variable that ends with an “s” is a school level variable, any variable ending in “d” is a district level variable.

The test score variables are “scale scores” as opposed to “raw scores”. The reading and mathematics scores range from 0 to 500 with the average score for 1993 set to 250. The writing scores range from 6 to 32. Scale scores are developed from a Rasch latent variable model and set to be comparable to the results from previous years. The Rasch model estimates from answers to each question for each test-taker one parameter for each question reflecting how easy that question is and an ability parameter for each student. (Rasch, 1966) In the first year of testing, Illinois transformed the student ability scores giving them a mean of 250 and a standard deviation of 100.<sup>5</sup> In subsequent years, similar conversions are made. However, to assure comparability across years, several questions from earlier exams are asked on the current exam. In the Rasch model, the “easiness” parameters for these repeated questions are constrained to be those

---

<sup>5</sup>The typical software used for estimating the Rasch model imposes a mean of 0 and a standard deviation of 1 on the individual ability parameters.

obtained on the earlier data.

#### IV) Results

This section of the paper reports the results of the estimations of the models described above. Table 3 shows the results of estimating equation 1 on the 1995 data for each of the three tests, reading, math, and writing. The results are typical in that average class size is not statistically significant in any of the equations, though it does have the expected sign in both the reading and the math estimations. Results for 1996 and 1997 are very similar to these. Of the nine estimates of the effect of class size on test scores, just five have the correct sign and only two of those are statistically significant at conventional levels.<sup>6</sup>

The estimates in Table 3 also indicate the commonly found importance of income, attendance and mobility, the enrollment in the school and the racial composition of the student body. These variables all have the intuitive signs. For example, the higher the proportion of the student body that is in a family which receives income support from the government, the lower are achievement test scores. The greater the proportion of students who attend school everyday, the higher are test scores and the higher is the proportion of the student body that moves in or out of the school during a school year the lower are test scores. High parental involvement also is positively and statistically significantly related to test scores.

Low proficiency in English amongst the student body would seem intuitively to suggest

---

<sup>6</sup>However, if average class size enters the equation quadratically, the linear and quadratic terms have opposite sign, each is generally statistically significant at the 10% level or better, and the quadratic term typically has the higher absolute t. Moreover, evaluated at the mean of the average class size, the marginal effect of class size is negative, as one would expect.



lower test scores. Yet that variable is found to be positive and statistically significant. However, only bilingual students who have participated in a “state-approved Transitional Bilingual Education (TBE) program or Transitional Program of Instruction (TPI)” for at least three years are required to take the IGAP examinations. Students in such programs for less than three years sit for the exams only if local school personnel judge their English proficiency sufficient for them to understand the tests. Consequently, low English proficiency students are not included in the sample.

None of this explains the strong positive association between low English proficiency and IGAP scores. One possibility is that this variable is a strong proxy for motivation and effort. The efforts of students from Asia to succeed academically in the United States despite language barriers are well publicized.

Teacher characteristics such as experience, education and race are inconsistent in sign and significance across equations. In Table 3, the percent of teachers with bachelors degrees and the percent with more than a masters degree each have positive but insignificant coefficients in the reading and math equations, but negative and insignificant coefficients in the writing equation. But for the 1996 and 1997 data these variables frequently have negative and significant coefficients. Teacher experience carries negative coefficients about half the time but these are not statistically significant. The percent of teachers that are white has negative coefficients throughout and is almost always significantly different from zero. Unlike the previously discussed variables, however, these are measured at the district level rather than at the school level. As they are possibly quite poor measures of the characteristics of the teachers of third grade in any given

school, the coefficient estimates are highly suspect.<sup>7</sup>

For each subject matter and for each year, tests were performed of both the null hypothesis that all instructional time variables have zero coefficients and the null that all instructional times and instructional times interacted with average class size have zero coefficients. Tables of the coefficient estimates are available upon request. The key finding here is that these variables clearly belong in the regression equation. Only for the writing test is it ever impossible to reject the null at standard probability levels. Even for the writing test, the null must be rejected for either formulation, instructional time or instructional time and interaction, for 1997, and for the exclusion of instructional times for 1996. Discussion of the coefficient estimates is put off until all structural form issues are addressed.

Table 4 presents estimates when the years are pooled. Pooling was tested in two ways for each test. First, the hypothesis that both the slopes and the intercepts varied by year was tested against the hypothesis that all were the same for every year. The hypothesis of identical slope and intercept coefficients was rejected. However, restricting the slope coefficients to be the same across years but allowing year-specific intercepts gave a different result. In this case, one could not reject the restrictions. In other words, the effects of the inputs are the same across years, but the mean of the test scores is not after controlling for systematic effects of inputs. Note the strong significance of at least one year specific dummy variable in every equation. Moreover, though the results are not reported in the table, pooling cannot be rejected for models including

---

<sup>7</sup>Interestingly, this analysis stands on its head Hanushek's critique that those studies which find significance school resources tend to use more aggregated data.

instructional times or instructional times and average class size interacted with those times.<sup>8</sup>

Pooling the three years does not improve the outlook with respect to the average class size and teacher characteristics variables, however. These variables continue frequently to be of the wrong sign and their statistical significance is not stable. The characteristics of the student body are consistent in sign and significance between the individual year regressions and the pooled regressions.

Table 5 provides the estimates from the pooled data when school specific intercepts are introduced. The year specific intercepts remain large and statistically significant. Among the remaining variables there are impressive differences resulting from inclusion of the school specific intercepts.

The sign of the average class size now accords with intuition and the statistical significance of this variable is undeniable ( the t-statistic nearest to indicating insignificance is - 3.46). The size and significance of the student characteristics are dramatically affected by inclusion of the school specific intercepts, though the signs of the variables are unaffected. The most dramatic impact is on the low English proficiency variable which is no longer statistically significant at any conventional level.

On the other hand, the teacher characteristic variables remain a puzzle. The teacher education variables become strongly significant but of the wrong sign in both the reading and math equations. Experience is statistically significant in only the math equation, at the 10% level, but carries the wrong sign. And the percent of district teachers that are white becomes positive in every equation, though significant in only the math equation and that only at the 10%

---

<sup>8</sup>The tables of coefficient estimates are available upon request.

level.

The null hypothesis that all the school specific intercepts are equal is rejected easily for every equation. The adjusted R squared rises from .69 to .86 in the reading equation, from .57 to .85 in the math equation and from .42 to .76 in the writing equation. In other words, the model fits substantially better, explaining at a minimum 17%, and up to 34%, more of the variation in test scores after the school specific intercepts are included than before they were introduced.<sup>9</sup>

Table 6 returns to the standard style of the education production function, not including the instructional times or the interactions between times and class size, but introduces the square of the average class size.<sup>10</sup> Ferguson (1991) and Ferguson and Ladd (1996) contend that class size enters non-linearly into the production function. Their approach is to use dummy variables for different ranges of class size. Here non-linearity is captured by both average class size and average class size squared in the production function. Note that in each of the three equations both variables are significant at least at the 10% level. Moreover, five of the six coefficients are significant at the 5% level or even better.

Interestingly, in each case, the marginal effect of average class size at very small sizes is positive. However, each marginal effect turns negative. This result suggests that economies of scale in class size exist for small class sizes but that eventually all such economies are exhausted. The class size at which the marginal effect of class size turns negative is between 18.8, for

---

<sup>9</sup>In Illinois, school districts with third grades can be either Unit or Elementary districts. Tests for pooling of the district types indicate that pooling is appropriate. Chicago schools comprise roughly one-third of the schools from Unit districts. It is not possible to test for pooling with Chicago schools and all others because there is no variation in the Chicago schools' teacher variables which are measured district wide.

<sup>10</sup>Table 6 is based upon the pooled sample over the three years as is obvious from the inclusion of the year specific dummy variables.

reading, and 23.5, for writing. The turning point is at 19.1 students in the math equation.

Average class size is about 23.1. Consequently, for most classrooms, the effect of increasing the number of students is to reduce the average test score.

Consider Table 8, however. In this table, the results are for the same model as in Table 6, except that the school specific fixed effects are introduced. Note that the class size variables no longer matter in the writing equation. But more interestingly, the sign of the class size and class size squared coefficients flip. Now as class size rises from 1 student on, the average test scores in both reading and mathematics decline, and continue to do so through class sizes of 28 to 30. Note that these turning points occur at from 5 to 7 more students than in the average class in the sample.

At the mean of the average class size, the marginal effect of an additional student is -.36 in reading, -.54 in math and -.02 in writing. One standard deviation above and below the mean each of these marginal effects remains negative. Relative to the mean of the test scores, these effects are very small. For example, raising the average class size from 23 to 24 results in a reduction in the average test score from 245.1 to 244.74. These effects hardly seem worth worrying about. Even raising the class size by a full standard deviation, from 23 to almost 28 students per class has no appreciable effect, less than one point, on the average score.

Other results in Table 6 look very much like those of Table 4. The school, district and student characteristics variables and the year specific intercepts all have essentially the same size and the same level of significance between the two models. The adjusted R squared is unaltered. The only effect of the nonlinearity in average class size is to make class size and its square statistically significant. But over the relevant range, average class size has very little practical

impact on the IGAP scores.

Introduction of the instructional times and their interactions with the average class size cannot be rejected. Table 7 reports the estimates of the production function when these variables are included. Note that in every equation at least three of these variables is individually significant at the 5% level. They seem also to indicate interesting things about the nature of education. For example, more time spent in mathematics instruction leads to improved reading scores. However, more instruction in language arts has no impact on math scores. Social studies instruction has negative consequences for math scores, but positive implications for reading and writing scores. Science instruction reverses the pattern of social studies, reducing reading and writing scores and raising those on the math test.

Table 9 repeats the analysis of Table 7 with the addition of the school fixed effects. (Alternatively, it extends the model of Table 8 to include the instructional time and interaction variables.) The first point to note is that after introduction of the school fixed effects the number of the time variables or time and class size interactions variables that are individually significant drops dramatically. For example, in Table 7 of the 24 coefficients across the three equations 13 are individually significant at the 5% level and one at the 10% level; in Table 9 only five are significant at either the 5 or the 10% levels. Indeed, for this model, the hypothesis of joint insignificance of these variables cannot be rejected at standard levels in the reading equation and is rejected for the math equation only at the 10% level. The hypothesis is rejected at the 5% level in the writing equation.

Just as was the case in the standard model when the school fixed effects were introduced, other parameter estimates were dramatically affected in size or significance or both. For example,

in Table 7, the low English proficiency variable is highly significant in every equation. Yet in Table 9 it is never even close to significance. Low income status of the school drops from extremely highly significant, with absolute t-statistics of 20 or more, to significant at the 5% level in only the reading equation. The coefficients fall in size by several times; in the reading equation from  $-.764$  to  $-.146$ , for example. Teacher characteristics variables are still of the wrong signs and hugely significant.

Unlike the class size variables, the effects of instructional times seem more substantial. For example, suppose that time spent on math instruction were raised by ten minutes a week. That would raise the average math score almost 1.5 points. It would also raise reading scores by half a point, which is more than enough to compensate for raising the average class size by one student.

Suppose now that we examine the possible interactions of average class size with the teacher characteristics. Tables 10 and 11 present the results when these variables are introduced to the model without and with the instruction times and class size-instruction time interactions, respectively. As is clear, adding the interactions to the standard model with the fixed effects reveals that those interactions belong. However, in the reading and math equations, inclusion of these variables indicates that the interactions with instructional times and the instructional time variables do not belong. Only for the writing equation can one reject the null hypothesis of joint insignificance of the set of interactions between average class size and instructional time variables and the levels of the time variables.

Finally, Table 12 reports the effects of adding interactions between average class size and the characteristics of the school and its student body. F-tests indicate that inclusion of these

additional interactions is warranted in the case of the writing test but not warranted in the reading or math test equations.

The discussion so far has been stilted toward “statistical significance” of variables. From a policy perspective, one would wish to know if the issues raised in this paper have any practical relevance. For example, much of the literature indicates that more and better trained teachers are essentially irrelevant, at least until the internal workings of the schools are straightened out. The analysis of this paper indicates that average class size does matter, and does so in rather complicated ways. In the remainder of this section, the effects of raising class sizes are assessed. Unlike the typical completely linear noninteracted models prevalent in the literature in which the marginal effect of increasing class size is simply the regression coefficient, the marginal effects from this analysis must be computed from the parameter estimates and the sample data.

The computations reveal something quite interesting. If one looks at the derivative of the test score equation with respect to class size, one obtains an equation like (3). Calculating this derivative, one obtains a negative number very near zero in each and every equation. Indeed, the size of this derivative is rather remarkably stable across models from Tables 7 through Table 12. However, if one examines the data for the later tables, 10 through 12, which the data supports against the earlier simpler models, one finds an important nuance. The effects of raising average class size if all the variables with which class size is interacted were set to zero is substantial and negative. It is only because of the countervailing effects of well trained teachers that some increases in average class size do not have disastrous consequences.

Table 13 shows the marginal effects of raising average class size, calculated at the sample means. The effect if all the interaction variables are set equal to zero is in the first column. Note



that the fall in average scores is 25% for both reading and writing and 17% for math. Once all the countervailing measures of teacher quality, instructional times and student and school characteristics are accounted for, the percentage declines from raising the average class size are less than 1% for both reading and math and about 2% for writing.

## Conclusion

This paper has criticized the existing education production function literature for not paying enough attention to the complete set of insights that production theory affords. Specifically, the theory suggests that there should be a high degree of interaction between the inputs into the production process. Tests of this hypothesis indicate that in fact such interactions are statistically significant. Moreover, computation of marginal impacts of average class size and of other inputs once interactions are included imply that schools can and do effectively use greater quantities of resources presumed to have beneficial effects to counteract greater quantities of those inputs thought to have harmful effects.

Additionally, the analysis has used panel data to allow for school specific fixed effects. The inclusion of these effects is found to be warranted and to have large impacts on the estimated effects of other variables. Introducing these effects into the standard model of an education production function was sufficient to bring average class size into statistical significance, for example.

Finally, the panel data enabled a test of the stability of test scores from year to year. In this regard, the analysis clearly indicates that it may be unwise to examine test score reports for simple changes to draw conclusions about school performance. Year specific dummy variables

are routinely significant. Consequently, one cannot assume that a school has performed better based on the knowledge that its scores rose from one year to the next. The effect could be either better performance by the school in educating the children or simply that the test was easier. Policy makers would be wise therefore not to allot resources based on year to year differences in these test scores.

## Bibliography

- Betts, Julian. "Which Types of Public School Spending are Most Effective? New Evidence on the School Quality Debate," ...
- \_\_\_\_\_. "The Role of Homework in Improving School Quality," ...
- \_\_\_\_\_. "Do Grading Standards Affect the Incentive to Learn?"...
- \_\_\_\_\_ and Jamie L. Shkolnik. 1996. "The Effects of Ability Tracking on Student Math Achievement and Resource Allocation in Secondary Schools". UCSD Discussion Paper 96-25.
- Clotfelter, Charles T. and Helen F. Ladd. 1996. "Recognizing and Rewarding Success in Public Schools," in Holding Schools Accountable: Performance-Based Reform in Education, edited by Helen F. Ladd. Washington, D.C.: Brookings Institution Press.
- Cooper, Harris. 1989. Homework. New York, NY: Longman.
- Ehrenberg, Ronald G. And Dominic J. Brewer. 1994. "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," Economics of Education Review, 31(1):1-17.
- Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters," Harvard Journal on Legislation, 28 (Summer):465-97.
- \_\_\_\_\_ and Helen F. Ladd. 1996. "How and Why Money Matters: An Analysis of Alabama Schools," in Holding Schools Accountable: Performance-Based Reform in Education, edited by Helen F. Ladd. Washington, D.C.: Brookings Institution Press.
- Fredrick, Wayne C. And Herbert J. Walberg. 1980. "Learning as a Function of Time," Journal of Educational Research, 73: 183-194.
- Hanushek, Eric A. 1996. "School Resources and Student Performance," in Does Money Matter?: The Effect of School Resources on Student Achievement and Adult Success, edited by Gary Burtless. Washington, D. C.: Brookings Institution Press.
- \_\_\_\_\_, Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and the Estimated Effects of School Resources," Review of Economics and Statistics, 78(4):611-627.
- Hedges, Larry V. And Rob Greenwald. 1996. "Have Times Changed? The Relation Between School Resources and Student Performance," in Does Money Matter?: The Effect of School Resources on Student Achievement and Adult Success, edited by Gary Burtless.

Washington, D. C.: Brookings Institution Press.

- Keith, Timothy Z., .... 1986. "Parental Involvement, Homework, and TV Time: Direct and Indirect Effects on High School Achievement," Journal of Educational Psychology, 78:373-380.
- Krueger, Alan B. 1997. "Experimental Estimates of Education Production Functions". Princeton University, Industrial Relations Section, working paper #379.
- Ladd, Helen F., Becky L. Roselius and Randall P. Walsh. 1997. "Using Student Test Scores to Measure the Effectiveness of Schools", paper presented at the November, 1997 APPAM Conference.
- Ludwig, Jens, and Laurie J. Bassi. 1997. "School Spending and Student Achievement: New Evidence from Longitudinal Data", paper presented at the November, 1997, APPAM Conference.
- Rasch, Georg. 1966. "An Individualistic Approach to Item Analysis," in Readings in Mathematical Social Science, edited by Paul F. Lazarsfeld and Neil W. Henry. Cambridge, Massachusetts: The M.I.T. Press.
- Somers, Julie. 1997. "Reconciling Rational Household Behavior with Estimated Education Production-Functions", paper presented at the November, 1997, APPAM Conference.
- Stevenson, Harold W. And James W. Stigler. 1992. The Learning Gap: Why our Schools are Failing and What We Can Learn from Japanese and Chinese Education, New York, NY: Summit Books.
- Summers, Anita A. And Barbara L. Wolfe. 1977. "Do Schools Make a Difference?," American Economic Review, 67(4):639-652.
- Walberg, Herbert J., Barry J. Fraser, and Wayne W. Welch. 1986. "A Test of a Model of Educational Productivity among Senior High School Students," Journal of Educational Research, 79:133-139.

Table 1: Variable Definitions

PCTWHTSC	The percent of the students enrolled in the school that are white
ENROLLSC	Number of students enrolled in the school
LEPSCHL	The percent of the students eligible for bilingual education
LINCOMES	The percent of the students in the school from families receiving public assistance, living in institutions for neglected or abused children, being supported in foster homes with public funds, or eligible for free or reduced price lunches.
PARNVLVS	The percent of the children in the school whose parents made at least one contact with the student's teachers during the school year
ATNDRTS	The percent of the students who attended school every day.
MOBRATES	The percent of the students who enroll in or leave a school during the school year.
ACS3S	Total enrollment for third grade divided by the number of third grade classes on the first school day of May
WHTTCHRD	The percent of the district's teachers who are white
BACHD	The percent of the district's teachers with a Bachelor's degree
MAPLUSD	The percent of the district's teachers with a Master's degree or more
TCHEXPD	Average years of teaching experience for teachers in the district
MPDENG3S	Average minutes of English instruction during a 5-day school week in the third grade classes in the school (includes all language arts)
MPDMTH3S	Average minutes of math instruction during a 5-day school week in the third grade classes in the school
MPDSC3S	Average minutes of science instruction during a 5-day school week in the third grade classes in the school
MPDSS3S	Average minutes of social studies instruction during a 5-day school week in the third grade classes in the school
ACMPDE	Interaction between ACS3S and MPDENG3S
ACMPDM	Interaction between ACS3S and MPDMTH3S
ACMPDSS	Interaction between ACS3S and MPDSS3S

Table 1: Variable Definitions

ACMPDSC	Interaction between ACS3S and MPDSC3S
YR95	Dummy variable taking on 1 if the observation is from 1995, zero otherwise
YR96	Dummy variable taking on 1 if the observation is from 1996, zero otherwise
READING	The average score on the reading portion of the IGAP exam among third graders in the school; possible scores range from 0 to 500
MATH	The average score on the math portion of the IGAP exam among third graders in the school; possible scores range from 0 to 500
WRITING	The average score on the writing portion of the IGAP exam among third graders in the school; possible scores range from 6 to 32

Table 2: Variable Descriptive Statistics								
	Full Sample		1995		1996		1997	
	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.
PCTWHTSC	66.24	36.77	67.04	36.51	66.14	36.85	65.55	36.93
ENROLLSC	443.75	237.28	437.82	230.25	444.83	237.57	448.62	243.83
LEPSCHL	5.59	12.11	5.33	11.62	5.60	12.10	5.82	12.60
LINCOMES	38.59	31.26	37.61	30.77	38.73	31.37	39.43	31.61
PARNVLVS	96.43	9.61	95.93	11.04	96.64	8.89	96.71	8.71
ATNDRTS	95.05	1.74	95.04	1.81	94.94	1.84	95.18	1.56
MOBRATES	20.36	13.92	20.89	14.44	20.23	13.47	19.95	13.83
ACS3S	23.06	4.66	23.25	4.76	22.71	4.54	23.21	4.65
WHTTCHRD	85.04	22.53	85.08	22.63	85.02	22.55	85.02	22.42
BACHD	60.05	15.06	60.89	14.64	60.37	14.95	58.88	15.50
MAPLUSD	39.77	15.04	38.96	14.63	39.44	14.93	40.93	15.47
TCHEXPD	14.27	2.03	13.98	1.97	14.23	2.01	14.59	2.07
MPDENG3S	146.95	24.55	147.11	24.02	147.01	23.96	146.72	25.63
MPDMTH3S	52.43	9.14	51.93	8.81	52.21	8.85	53.14	9.68
MPDSS3S	29.36	7.96	29.21	8.01	29.30	7.79	29.58	8.07
MPDSC3S	29.19	8.34	29.11	8.32	28.99	8.20	29.47	8.50
ACMPDE	3395.50	905.15	3425.80	908.27	3346.20	882.10	3414.50	922.97
ACMPDM	1206.90	314.83	1205.50	312.64	1186.10	313.76	1229.10	316.73
ACMPDSS	673.20	221.94	675.91	229.37	661.97	217.88	681.73	217.99
ACMPDSC	668.42	226.12	672.56	229.80	654.51	223.16	678.20	224.75
YR95	0.33	0.47						
YR96	0.33	0.47						
Reading	245.1	55.3	244.2	55.2	247.4	55.7	243.8	54.8
Math	281.3	60.3	272.5	61.3	285.3	60.2	286	58.5
Writing	18.1	2.11	18.3	2.05	17.7	2.17	18.4	2.03
Observations			2271		2270		2266	

Table 3: Standard Education Production Functions						
	Reading 95		Math 95		Writing 95	
	Coeff.	t-statistic	Coeff.	t-statistic	Coeff.	t-statistic
Constant	-910.560	-2.779	-952.500	-2.568	-4.739	-0.359
PCTWHTSC	0.459	9.137	0.391	6.250	0.005	2.153
ENROLLSC	-0.008	-2.207	-0.011	-2.368	0.000	0.038
LEPSCHL	0.160	2.218	0.230	2.640	0.010	2.982
LINCOMES	-0.699	-13.350	-0.786	-12.141	-0.026	-10.628
PARNVLVS	0.214	2.974	0.048	0.537	0.010	3.179
ATNDRTS	7.493	10.190	8.104	8.999	0.240	7.771
MOBRATES	-0.222	-2.828	-0.235	-2.775	-0.004	-1.426
ACS3S	-0.081	-0.518	-0.260	-1.286	0.000	0.007
WHTTCHRD	-0.330	-4.229	-0.321	-3.166	-0.007	-2.110
BACHD	4.407	1.355	4.766	1.283	-0.001	-0.009
MAPLUSD	4.935	1.516	5.379	1.449	0.020	0.153
TCHEXPD	-0.408	-1.055	-0.312	-0.621	-0.006	-0.306
Adj. R sqrd.	0.68		0.58		0.38	



Table 4: Standard Education Production Functions Pooled Sample						
	Reading		Math		Writing	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
Constant	-430.650	-2.537	-134.580	-0.607	1.253	0.162
PCTWHTSC	0.447	16.114	0.412	12.049	0.007	5.109
ENROLLSC	-0.008	-3.928	-0.009	-3.678	0.000	0.130
LEPSCHL	0.131	3.235	0.200	3.998	0.008	4.238
LINCOMES	-0.760	-25.128	-0.776	-20.657	-0.028	-19.068
PARNVLVS	0.154	3.278	0.143	2.516	0.006	2.619
ATNDRTS	7.029	17.569	7.627	15.158	0.232	12.251
MOBRATES	-0.245	-5.812	-0.238	-4.526	-0.003	-1.520
ACS3S	-0.141	-1.522	-0.242	-1.990	0.000	0.054
WHTTCHRD	-0.299	-6.511	-0.318	-5.365	-0.009	-3.955
BACHD	0.101	0.060	-2.992	-1.356	-0.048	-0.630
MAPLUSD	0.584	0.347	-2.405	-1.090	-0.025	-0.328
TCHEXPD	-0.150	-0.677	-0.033	-0.117	-0.017	-1.476
YR95	0.510	0.555	-12.940	-10.931	-0.121	-2.551
YR96	5.122	5.564	1.136	0.960	-0.690	-14.374
Adj. R sqrd.	0.69		0.57		0.42	

Table 5: Standard Education Production Function with School Specific Intercepts						
	Reading		Math		Writing	
	Coef.	t-stat.	Coef.	t-stat.	Coef.	t-stat.
PCTWHTSC	0.476	3.037	0.536	3.025	0.011	1.397
ENROLLSC	-0.010	-1.176	-0.018	-1.851	0.000	0.758
LEPSCHL	0.106	0.625	0.058	0.303	0.013	1.492
LINCOMES	-0.146	-2.141	-0.060	-0.776	-0.005	-1.408
PARNVLVS	0.067	1.448	0.054	1.024	0.000	0.093
ATNDRTS	1.738	3.223	2.143	3.515	0.040	1.448
MOBRATES	-0.107	-2.119	-0.072	-1.264	0.003	0.972
ACS3S	-0.349	-3.665	-0.527	-4.890	-0.017	-3.432
WHTTCHRD	0.603	1.188	0.958	1.670	0.022	0.859
BACHD	-9.725	-2.172	-21.403	-4.229	0.125	0.548
MAPLUSD	-9.559	-2.131	-21.108	-4.164	0.131	0.573
TCHEXPD	-0.262	-0.368	-1.482	-1.841	0.001	0.021
YR95	0.519	0.618	-13.957	-14.722	-0.148	-3.455
YR96	3.726	5.357	-1.090	-1.386	-0.743	-20.994
Adj. R sqrd.	0.86		0.85		0.76	

Table 6: Standard Education Production Function with Nonlinear Class Size						
	Reading		Math		Writing	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
PCTWHTSC	0.449	17.084	0.416	12.340	0.007	5.257
ENROLLSC	-0.008	-4.202	-0.010	-3.924	-0.000	-0.278
LEPSCHL	0.136	3.534	0.207	4.201	0.009	4.269
LINCOMES	-0.762	-27.915	-0.779	-22.280	-0.028	-20.081
PARNVLVS	0.154	3.613	0.142	2.616	0.006	2.802
ATNDRTS	7.049	19.251	7.660	16.334	0.234	12.293
MOBRATES	-0.244	-6.375	-0.238	-4.836	-0.003	-1.452
ACS3S	0.827	1.759	1.338	2.221	0.094	3.854
ACSSQ	-0.022	-2.095	-0.035	-2.670	-0.002	-3.911
WHTTCHRD	-0.309	-7.375	-0.335	-6.244	-0.010	-4.400
BACHD	0.020	0.012	-3.124	-1.539	-0.056	-0.677
MAPLUSD	0.495	0.312	-2.551	-1.256	-0.034	-0.408
TCHEXPD	-0.170	-0.802	-0.066	-0.243	-0.019	-1.697
YR95	0.502	0.539	-12.953	-10.869	-0.122	-2.525
YR96	5.070	5.485	1.051	0.888	-0.695	-14.475
C	-433.183	-2.703	-138.725	-0.676	1.007	0.121
Adj. R sqrd.	0.69		0.57		0.42	

Table 7: Standard Education Production Function with Instruction Times and Class Size Interactions

	Reading		Math		Writing	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
PCTWHTSC	0.447	17.024	0.410	12.216	0.007	5.332
ENROLLSC	-0.008	-4.221	-0.010	-3.874	-0.000	-0.515
LEPSCHL	0.135	3.514	0.209	4.267	0.008	4.191
LINCOMES	-0.764	-27.954	-0.788	-22.546	-0.028	-19.802
PARNVLVS	0.142	3.335	0.130	2.396	0.006	2.664
ATNDRTS	6.986	19.105	7.600	16.261	0.234	12.301
MOBRATES	-0.255	-6.661	-0.256	-5.231	-0.004	-1.763
ACS3S	1.695	1.709	1.217	0.960	0.025	0.479
ACSSQ	-0.020	-1.871	-0.028	-2.026	-0.002	-3.281
WHTTCHR	-0.335	-7.885	-0.391	-7.210	-0.010	-4.543
BACHD	0.541	0.337	-2.667	-1.299	-0.032	-0.378
MAPLUSD	1.004	0.625	-2.113	-1.029	-0.010	-0.122
TCHEXP	-0.204	-0.956	-0.082	-0.300	-0.023	-2.047
YR95	0.582	0.625	-12.665	-10.651	-0.133	-2.748
YR96	5.191	5.621	1.321	1.119	-0.703	-14.633
MPDENG3S	0.107	1.375	0.090	0.904	-0.001	-0.169
MPDMTH3S	0.595	2.842	0.763	2.852	-0.005	-0.445
MPDSS3S	-1.037	-3.392	-1.242	-3.178	-0.056	-3.501
MPDSC3S	0.818	2.785	0.638	1.700	0.033	2.174
ACMPDE	-0.001	-0.413	-0.000	-0.020	0.000	0.982
ACMPDM	-0.020	-2.231	-0.018	-1.537	0.000	0.299
ACMPDSS	0.045	3.324	0.047	2.714	0.002	3.370
ACMPDSC	-0.035	-2.754	-0.025	-1.502	-0.002	-2.333
C	-513.038	-3.143	-200.642	-0.962	-0.071	-0.008
Adj. R sqrd.	0.69		0.58		0.43	

Table 8: Education Production Function with Nonlinear Class Size and School Fixed Effects

	Reading		Math		Writing	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
PCTWHTSC	0.481	3.072	0.542	3.060	0.011	1.396
ENROLLSC	-0.009	-1.111	-0.017	-1.786	0.000	0.757
LEPSCHL	0.106	0.628	0.059	0.306	0.013	1.492
LINCOMES	-0.145	-2.131	-0.059	-0.764	-0.005	-1.408
PARNVLVS	0.069	1.484	0.055	1.061	0.000	0.092
ATNDRTS	1.798	3.334	2.211	3.626	0.040	1.444
MOBRATES	-0.109	-2.158	-0.074	-1.303	0.003	0.973
ACS3S	-1.824	-3.272	-2.206	-3.501	-0.015	-0.525
ACSSQ	0.032	2.685	0.036	2.704	-0.000	-0.063
WHTTCHRD	0.622	1.227	0.980	1.709	0.022	0.858
BACHD	-10.270	-2.293	-22.023	-4.350	0.126	0.550
MAPLUSD	-10.114	-2.254	-21.741	-4.287	0.131	0.575
TCHEXPD	-0.256	-0.360	-1.475	-1.834	0.001	0.021
YR95	0.496	0.591	-13.983	-14.760	-0.147	-3.454
YR96	3.745	5.388	-1.068	-1.359	-0.743	-20.991
Adj. R sqrd.	0.86		0.85		0.76	

Table 9: Education Production Function with Instructional Times and Interactions with Class Size and School Fixed Effects						
	Reading		Math		Writing	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
PCTWHTSC	0.478	3.046	0.548	3.090	0.011	1.406
ENROLLSC	-0.009	-1.086	-0.017	-1.746	0.000	0.825
LEPSCHL	0.114	0.676	0.064	0.333	0.013	1.548
LINCOMES	-0.146	-2.140	-0.054	-0.705	-0.005	-1.494
PARNVLVS	0.071	1.536	0.053	1.020	0.001	0.219
ATNDRTS	1.760	3.259	2.209	3.621	0.037	1.356
MOBRATES	-0.108	-2.130	-0.073	-1.280	0.003	1.033
ACS3S	-0.086	-0.083	-0.286	-0.242	-0.031	-0.583
ACSSQ	0.028	2.360	0.033	2.445	-0.000	-0.298
WHTTCHR	0.672	1.317	1.014	1.760	0.023	0.890
BACHD	-9.820	-2.187	-21.841	-4.304	0.164	0.715
MAPLUSD	-9.677	-2.152	-21.567	-4.243	0.168	0.734
TCHEXP	-0.306	-0.430	-1.546	-1.922	0.002	0.045
YR95	0.550	0.654	-13.821	-14.542	-0.145	-3.382
YR96	3.800	5.440	-0.887	-1.123	-0.742	-20.856
MPDENG3S	0.111	1.276	0.093	0.946	-0.004	-0.890
MPDMTH3S	0.336	1.498	0.600	2.369	-0.009	-0.769
MPDSS3S	0.478	1.520	0.549	1.546	-0.003	-0.205
MPDSC3S	-0.046	-0.148	-0.240	-0.684	0.036	2.286
ACMPDE	-0.003	-0.898	-0.002	-0.524	0.000	1.075
ACMPDM	-0.012	-1.244	-0.019	-1.786	0.001	1.142
ACMPDSS	-0.012	-0.870	-0.026	-1.672	0.001	0.918
ACMPDSC	-0.004	-0.329	0.011	0.720	-0.002	-2.753
Adj. R sqrd.	0.86		0.85		0.76	

Table 10: Teacher Characteristics Interacted with Average Class Size with School Specific Effects

	Reading		Math		Writing	
	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
PCTWHTSC	0.459	2.925	0.516	2.914	0.010	1.275
ENROLLSC	-0.009	-1.061	-0.017	-1.773	0.000	0.815
LEPSCHL	0.113	0.672	0.072	0.376	0.013	1.562
LINCOMES	-0.147	-2.160	-0.064	-0.825	-0.005	-1.443
PARNVLVS	0.073	1.573	0.063	1.198	0.001	0.242
ATNDRTS	1.801	3.321	2.116	3.451	0.038	1.379
MOBRATES	-0.108	-2.147	-0.075	-1.310	0.003	0.977
ACS3S	-62.948	-1.615	-49.892	-1.133	-4.167	-2.099
ACSSQ	0.023	1.832	0.021	1.457	-0.001	-0.962
WHTTCHRD	0.963	1.814	1.418	2.365	0.047	1.746
BACHD	-23.406	-2.474	-32.647	-3.053	-0.781	-1.620
MAPLUSD	-23.694	-2.504	-32.843	-3.072	-0.793	-1.645
TCHEXPD	1.391	1.156	-0.266	-0.196	0.051	0.832
YR95	0.434	0.514	-13.886	-14.556	-0.148	-3.439
YR96	3.821	5.456	-0.831	-1.050	-0.736	-20.614
ACBACH	0.630	1.600	0.498	1.119	0.043	2.125
ACMA	0.651	1.650	0.520	1.166	0.043	2.164
ACEXP	-0.081	-1.686	-0.055	-1.022	-0.002	-0.983
ACTCHRD	-0.012	-2.049	-0.017	-2.497	-0.001	-3.063
Adj. R sqrd.	0.86		0.85		0.76	

Table 11: Instructional Times Interacted with Class Size and Teacher Characteristics with School Specific Effects

	Reading		Math		Writing	
	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
PCTWHTSC	0.458	2.918	0.525	2.961	0.011	1.318
ENROLLSC	-0.009	-1.052	-0.017	-1.758	0.000	0.893
LEPSCHL	0.121	0.715	0.077	0.404	0.014	1.625
LINCOMES	-0.147	-2.152	-0.059	-0.762	-0.005	-1.514
PARNVLVS	0.075	1.611	0.061	1.163	0.001	0.391
ATNDRTS	1.793	3.306	2.140	3.491	0.038	1.370
MOBRATES	-0.108	-2.128	-0.073	-1.284	0.003	1.046
ACS3S	-59.443	-1.507	-36.416	-0.817	-4.960	-2.470
ACSSQ	0.022	1.756	0.020	1.403	-0.001	-0.932
WHTTCHRD	0.954	1.789	1.371	2.275	0.049	1.821
BACHD	-22.424	-2.354	-29.898	-2.778	-0.907	-1.869
MAPLUSD	-22.717	-2.385	-30.096	-2.796	-0.916	-1.889
TCHEXPD	1.223	1.015	-0.412	-0.303	0.056	0.909
YR95	0.448	0.529	-13.739	-14.348	-0.149	-3.448
YR96	3.815	5.423	-0.696	-0.876	-0.737	-20.576
ACBACH	0.606	1.522	0.373	0.828	0.050	2.476
ACMA	0.626	1.571	0.394	0.874	0.051	2.507
ACEXP	-0.076	-1.581	-0.052	-0.958	-0.003	-1.083
ACTCHRD	-0.010	-1.659	-0.015	-2.152	-0.001	-3.243
MPDENG3S	0.125	1.420	0.093	0.931	-0.004	-0.940
MPDMTH3S	0.244	1.057	0.423	1.623	-0.017	-1.446
MPDSS3S	0.440	1.391	0.576	1.608	-0.005	-0.298
MPDSC3S	-0.139	-0.429	-0.495	-1.355	0.030	1.828
ACMPDE	-0.004	-1.053	-0.002	-0.510	0.000	1.133
ACMPDM	-0.008	-0.803	-0.012	-1.047	0.001	1.827
ACMPDSS	-0.010	-0.740	-0.027	-1.736	0.001	1.013
ACMPDSC	-0.001	-0.045	0.022	1.393	-0.002	-2.285
Adj. R sqrd.	0.86		0.85		0.76	



Table 12: Education Production Function with School Specific Effects and Class Size Interactions with all Variables

	Reading		Math		Writing	
	Parameter	t-stat.	Parameter	t-stat.	Parameter	t-stat.
PCTWHTSC	0.642	3.057	0.753	3.177	0.027	2.523
ENROLLSC	-0.006	-0.380	-0.010	-0.545	0.002	2.434
LEPSCHL	0.030	0.108	0.058	0.186	-0.003	-0.197
LINCOMES	-0.127	-0.779	-0.062	-0.338	-0.008	-0.980
PARNVLVS	0.093	0.490	0.598	2.805	0.007	0.774
ATNDRTS	-0.178	-0.096	-1.614	-0.770	-0.060	-0.630
MOBRATES	-0.246	-1.299	0.006	0.030	0.020	2.116
ACS3S	-61.941	-1.476	-25.365	-0.535	-4.527	-2.121
ACSSQ	0.023	1.717	0.023	1.572	-0.000	-0.324
WHTTCHRD	0.783	1.401	1.183	1.876	0.034	1.198
BACHD	-20.823	-2.114	-24.207	-2.177	-0.718	-1.434
MAPLUSD	-21.081	-2.139	-24.431	-2.196	-0.728	-1.452
TCHEXPD	1.025	0.822	-0.141	-0.100	0.060	0.953
YR95	0.473	0.557	-13.720	-14.313	-0.143	-3.303
YR96	3.825	5.429	-0.718	-0.902	-0.736	-20.540
ACBACH	0.548	1.325	0.135	0.288	0.042	2.009
ACMA	0.566	1.366	0.157	0.335	0.043	2.037
ACEXP	-0.069	-1.363	-0.067	-1.177	-0.003	-1.123
ACTCHRD	-0.003	-0.300	-0.007	-0.697	-0.000	-0.859
MPDENG3S	0.129	1.457	0.080	0.802	-0.005	-1.025
MPDMTH3S	0.243	1.049	0.384	1.468	-0.017	-1.414
MPDSS3S	0.430	1.355	0.592	1.652	-0.003	-0.175
MPDSC3S	-0.148	-0.458	-0.541	-1.479	0.028	1.682
ACMPDE	-0.004	-1.082	-0.002	-0.392	0.000	1.215
ACMPDM	-0.008	-0.801	-0.010	-0.925	0.001	1.767
ACMPDSS	-0.010	-0.712	-0.028	-1.785	0.001	0.907
ACMPDSC	-0.000	-0.015	0.024	1.522	-0.002	-2.151
ACWHTSC	-0.008	-1.331	-0.010	-1.474	-0.001	-2.277
ACENROLL	-0.000	-0.224	-0.000	-0.480	-0.000	-2.321
ACLEP	0.003	0.365	0.000	0.044	0.001	1.465
ACINC	-0.001	-0.118	-0.000	-0.002	0.000	0.354
ACPAR	-0.001	-0.098	-0.023	-2.585	-0.000	-0.691
ACATND	0.085	1.121	0.158	1.832	0.004	1.084
ACMOB	0.006	0.749	-0.004	-0.405	-0.001	-1.917
Adj. R sqrd.	0.86		0.85		0.76	

Table 13: Marginal Effects of Average Class Size		
	W/O Interactions	With Interactions
Reading (Coefficients from Table 10)	-61.88738	-0.340839
Math (Coefficients from Table 10)	-48.92361	-0.567975
Writing (Coefficients from Table 12)	-4.480886	-0.408519