96-16

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

THE ROLE OF HOMEWORK IN IMPROVING SCHOOL QUALITY

BY

JULIAN R. BETTS

DISCUSSION PAPER 96-16
MAY 1996

The Role of Homework in Improving School Quality

UCSD Discussion Paper 96-16

Julian R. Betts
Department of Economics
University of California, San Diego
La Jolla, CA 92093-0508
(619) 534-3369
e-mail: jbetts@ucsd.edu

February 1996

I thank UCSD for research support, and Jeff Grogger, Rick Hanushek and participants at the 1996 Winter Meetings of the Econometric Society for helpful suggestions. I also would like to thank Eric Johnson for his excellent research assistance. Postcript copies of this paper are available on the World Wide Web from http://weber.ucsd.edu/~jbetts.

ABSTRACT

The paper extends the school quality literature by examining the impact of homework on educational achievement. Theory predicts a positive effect of homework, up to a point. The theory is tested using the Longitudinal Study of American Youth. The amount of math homework assigned positively affects math achievement for virtually all students. An extra half hour of nightly homework between Grades 7 and 11 is predicted to boost math achievement by almost two grade equivalents. Additional homework is potentially cost-effective, since it is the homework assigned, rather than the amount graded, which is more has the larger influence on achievement.

The Role of Homework in Improving School Quality

1. Introduction

The positive link between earnings and years of education has become established as one of the most robust findings in the labor economics literature.

Over the last three decades, spurred by the Coleman Report (1966), researchers have extended the simple earnings:education model to more complex models in which the *quality* of schooling, and not simply its quantity, influences student achievement and subsequent earnings. This 'school quality' literature has by now shown that substantial differences in quality exist between schools. However, with a few exceptions, the vast majority of papers on this issue have found that traditional yardsticks of school quality, such as teachers' credentials and class size, can explain only a small portion of the observed quality variations between schools.[1]

The literature on education production functions focuses on the role of financial inputs such as overall spending per pupil and class size. While this approach is useful, it has the drawback of treating students as inanimate intermediate inputs to whom value is 'added'. The goal of the present paper is to extend this literature by treating students not as mere inputs but as utility-maximizing agents whose effort level responds to the perceived rewards.

[1] For reviews of the link between school spending and test scores, see Hanushek (1986). For a review of the literature on the impact of school spending on earnings and educational attainment, see Betts (forthcoming) and Card and Krueger (1994).

The paper develops such a model and then tests it. Specifically, the paper tests the hypothesis that improvements in student achievement from one year to the next depend positively on the amount of homework which teachers assign. The model predicts that homework standards have a positive influence on student effort. However, the relationship is not monotonic because beyond a certain point a heavy homework load may cause weaker students to exert no effort at all. This non-linearity will be explicitly modeled.

Cooper (1989) provides a detailed review of earlier research on the link between homework and student achievement. He cites a number of experiments which suggest a positive link. However, the sample sizes in these studies are very small (39 to 400 students) and the studies examined only one to eight schools each. A larger literature examines the correlation between achievement and time spent on homework in a non-experimental cross-sectional framework. Cooper reports the results of 11 studies which model student achievement as a function of homework while controlling for background variables. Although most of the studies indicate a positive link between homework and achievement, in most cases the nature of the data limits the conclusions that can be drawn. Several studies involved small samples which are not representative national samples. Others used national samples but did not control well for prior achievement of the student, thus increasing the risk of omitted variable bias. Two notable exceptions are the work of Keith et al. (1986) and Walberg et al. (1986), which use High School and Beyond and the National Assessment in Science, respectively, to establish a correlation between student test scores and the amount of homework which the student reported doing per time period.

Unfortunately, these studies, like the vast majority of the literature, are forced to use a student report on hours of homework done per week. This is not a policy variable which a school administrator or teacher can directly control. In particular, there is a risk that most of the variation in homework performed by students in a school reflects unmeasured differences in student ability or attitudes. Another typical problem in the literature is that achievement in a given subject is regressed on homework performed in *all* subjects. The ideal measure of homework would be the amount of *homework assigned* by the student's teacher in the given subject.

The goal of this paper is to develop an explicit model of student achievement in which students respond to the homework standards imposed by their teachers, and then to test the model using a new dataset which transcends four of the problems outlined above. First, the dataset -- the Longitudinal Study of American Youth -- is large and nationally representative. Second, the data improve on much of the earlier work in that explicit controls are included for students' background and prior achievement. Third, and most important of all, the study uses reports by the student's math teacher on the *hours of homework assigned per week*. This measure of homework is arguably more useful than hours of homework performed by each student, due to the possibility that time spent on homework merely reflects unobserved variations in student motivation and ability. A fourth advantage of the new dataset is that the homework question is specific to the subject for which test scores are available. In addition to these data improvements, the paper also makes use of the panel structure of the data to test whether the results are robust to unobserved heterogeneity across students.

Section 2 develops an explicit behavioral model of learning. Section 3 describes the data, and Section 4 outlines the results.

2. Theoretical Model

This section develops a hybrid human capital/signaling model of education in which attendance at school adds to productivity. Firms can observe the standard which each student meets. In particular, they can identify those who graduate from high school and those who do not. (Firms can also observe whether the student has failed a grade.) This information signals to firms the expected productivity of each worker. Thus educational achievement also acts as a signal.

Several antecedents to this model of educational standards exist. See for example Kang (1985), Becker and Rosen (1990) and Costrell (1994).

The most closely related model to the one presented below is that of Costrell (1994). The principal difference between the two models is that the Costrell model assumes that all students are identical in productivity at zero effort levels, but that they differ in preferences. The model below assumes that students differ in ability. Unlike most earlier models of educational standards, apart from Costrell (1994), this model endogenizes wages as a function of educational standards. This is an important extension since it uses a standard utility function depending on wages and leisure to explain why students would care about their performance in school.

Each student has an initial level of achievement (or ability) of a, where a is a random variable distributed on the real line with probability density and cumulative distribution functions of f(a) and F(a) respectively. Student i maximizes utility, which depends on leisure (L_i) and lifetime earnings (w_i):

(2-1) $$U = U(L_i, w_i) \text{ subject to } L_i \in [0, \bar{L}]$$

Productivity in the labor market is proportional to educational achievement, π_i. Productivity of a worker of ability a_i who has chosen level a L_i of leisure is given by

(2-2) $$\alpha\pi_i = \alpha g(\bar{L} - L_i, a_i), \text{ s.t. } g_1 > 0,\ g_2 > 0,\ g_{11} < 0,\ g_{22} < 0,\ g_{12} > 0,\ \alpha > 0.$$

The school sets a standard for the minimum level of achievement necessary to pass, which is denoted π_s. This standard is determined by the workload assigned to students. The amount of homework which the teacher assigns is the main determinant of overall workload.

Firms cannot observe the productivity of workers directly. But firms do observe whether the student has met the standard imposed by the school, and thus infer his expected productivity. Perfect competition in the labor market ensures that workers are paid their expected product conditional on whether they have met the standard. Workers belong either to group 1, which consists of students who have met the standard π_s, or group 2, which consists of workers who have not met the standard.

Conditional on reaching the achievement level needed to belong to group j, j=1,2, the worker optimizes by choosing the maximum level of leisure possible. Thus workers in group 1, who meet the standard,

(2-3) $$\pi_i = \pi_s$$

Similarly, for workers in group 2, who do not meet the homework standard, $L_i = \bar{L}$.

<u>Result</u> If some students choose to reach the homework standard and others do not, then there exists an ability level a* such that workers choose group (1/2) as a ⋛ a*.

<u>Proof</u>

This result follows directly from the observation that utility is independent of ability for workers in group 2, and increasing in ability for those in group 1.

We can now calculate the expected productivity, and hence the wage, for workers in each group:

$$(2-4) \qquad w_i^1 = \alpha E(\pi_i | \pi_i = \pi_s) = \alpha \pi_s$$

and for workers who do not meet the standard,

$$(2-5) \qquad w_i^2 = \alpha E(\pi_i | \pi_i < \pi_s) = \\ = \frac{\alpha \int_{-\infty}^{a^*} g(0,a) f(a) da}{F(a^*)}$$

where all workers optimize by setting $L_i = \bar{L}$.

A secondary and interesting question concerns what happens when schools raise the homework standard π_s. For students who continue to meet the standard under the new regime, their achievement rises as does their wage:

$$(2-6) \qquad \left. \left(\frac{dw_i^1}{d\pi_s} \right) \right|_{a > a^*} = \alpha$$

Students of lower ability, in group 2, will continue to exert the minimum effort possible.

The most important implication of this model for the empirical work to follow is that a fraction of students at the bottom tail of the achievement distribution will not respond to changes in educational standards.

It can be shown that the change in the proportion of students who fail to meet the standard, F(a*), after an increase in the standard, is indeterminate in sign. But in the more likely (and more intuitive case), a student who was originally on the margin between meeting and not meeting the standard finds that the extra effort required to achieve the new higher standard is not worth it, and so joins the less able students in group 2.

Figure 1 illustrates the distribution of achievement π with respect to ability a. Students in the top group all exactly meet the standard, while among students in the bottom group educational achievement rises with ability since even with zero effort more able students learn more than less able students. The likely effect of an increase in the educational standard to π_s' is shown by the dotted line. The higher standard increases achievement for the top group of students, while causing a small proportion of 'borderline' students to fail to meet the new standard. The educational achievement of the bottom group of students is unaffected by the change.

Note that corner solutions are possible in this model. The standard may be set so high that no students meet it. Perhaps more plausibly, the homework standard may be set so low that all students choose to respond to the standard.

3. Data

The model is tested using the *Longitudinal Study of American Youth* (LSAY). The LSAY consists of a panel survey of 3116 students who were in Grade 7 and 2829 students who were in Grade 10 in fall 1987, which marks the first year of the survey. I use the first five years of the panel, as this is the extent of the data currently available. Students represent a random sample of approximately 60 students each from a sample of 52 high schools and 52 'feeder' middle schools. The schools were selected as a random sample stratified by 4 regions and 3 'community types' (urban, suburban and rural). Students filled out detailed surveys twice a year while in secondary school. They also took standardized tests in science and math each fall. The tests involved questions developed for the National Assessment of Educational Progress, and were structured to allow inter-year comparability. The questions were designed so that students in any grade were unlikely to score full marks, thus avoiding 'ceiling effects'. [2]

The students' parents answered questions once a year, as did the teachers of the students in math, science, and various other classes. Most crucially, the teacher reports the average weekly hours of homework which he or she assigned to the class. These surveys are supplemented by one-time surveys of the principal of each school (conducted in the 1989-90 school year) and of the science and math teachers of each student. The one-time teacher surveys were given in various waves between 1988 and 1991. Since one of the questions which this survey asks is the teacher's years of full-time teaching

[2] For a fuller description of the LSAY, see Miller et al. (1992).

experience, a teacher's experience in a given school year is calculated based on the assumption that the teacher has worked full-time in all years in the survey.

This paper uses a data-set obtained by merging each student's information with the relevant data from the one-time surveys of math and science teachers and principals, and the annual surveys of classroom teachers and parents. In a small percentage of cases, a student will have had two different math teachers in a given year. In this case, the average of each teacher and classroom trait is taken.

Both the younger and older cohorts of students are used in the analysis.

The model will be tested by examining the determinants of students' scores in standardized tests of math administered each fall.

4. Empirical Implementation and Results

The econometric approach is to model math test scores by extending a value-added education production function which treats homework as an additional measure of school inputs.[3] These models, which have become the preferred statistical method for modeling school quality, regress math test scores, S, in one period on a vector of personal, family and peer background variables F, a vector of school resources such as class size, R, and the student's test score from the previous period. (The lagged test score is added in order to control for the unobserved impact of schooling in earlier years of the student's life. The coefficient on this variable can be interpreted as one minus the rate of depreciation of prior human capital.) Thus for student i in grade g the basic model is:

[3] For a review of the existing literature on the impact of school and teacher traits on students' test scores, see Betts, 1995.

(4-1) $S_{ig} = \alpha + \beta S_{i,g-1} + \Gamma F_{ig} + \theta R_{ig} + \delta HW_{ig} + \varepsilon_{ig}$

where HW_{ig} is the amount of homework assigned. The theoretical model hypothesizes that $\delta > 0$. [4]

The key measures of school and classroom traits, which appear in the tables, are the weekly hours of homework assigned by the student's math teacher, the years of full-time teaching experience of the student's math teacher and its square, dummies for whether the math teacher holds an Associate degree, a Master's degree or a Ph.D., the number of students in the math class, overall enrollment in the school, and the percentage of students in the school who were eligible for full Federal lunch assistance.

The other regressors not shown in the tables are a lagged math test score, a constant, dummies for students in grades 8 through 11, 8 dummies indicating mother's and father's level of educational attainment, an index variable created by the administrators of the LSAY which proxies the family's socioeconomic status on a -3 to +3 scale,[5] dummies for students who are black, Hispanic, Asian or native American, dummies for whether the school is in a suburban or rural area, dummies for three of four national regions, and the percentage of the student body which is black, Hispanic or Asian. Together, these regressors provide a fairly rich set of controls for family background, characteristics of the student body and traits of the school, math teacher, and math class. (See Table A-1 in the Appendix for a list of means and standard deviations of the variables.)

The results appear in Table 1.

---

[4] The math tests were administered each fall. Thus for a given grade level g, the test score from the following autumn, S_{ig}, are regressed on classroom inputs from the current school year, and the test score obtained at the start of that school year, $S_{i,g-1}$.

[5] This index was calculated by the administrators of the LSAY as an average of variables indicating parents' level of education, parents' occupational status, and an index proxying for household possessions.

Column 1 shows the results for a basic test score equation without homework as a regressor. In the standard value-added models without homework, most of the measures of school inputs are statistically insignificant. There are three school variables which are significant at 10% or less: teachers' experience, the dummy for whether the teacher has a Master's degree, and the overall size of the student body. The first of these effects suggests that more experienced teachers are less effective in the classroom, although this effect diminishes as experience rises.[6] Teachers with Master's degrees are associated with larger gains in test scores. Finally, there is evidence of a weak diseconomy of scale in that smaller schools appear to be slightly more conducive to faster learning in mathematics.

Column 2 adds the hours of homework assigned. 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.

To test for the possibility that there are diminishing returns to the amount of homework assigned, column 3 uses a quadratic specification. In this model, the homework squared term is insignificant, suggesting that in the range of hours of homework assigned in the sample, there are approximately constant returns. [7]

Outliers in test scores are a concern in the education production function literature. If a student performs poorly on a test one year, due to illness or some other related factor, then the data will indicate a large drop in achievement followed by a large rise in the

---

[6] It should be noted that the impact of teacher experience on test scores does appear to be non-linear, as modeled. When the regressions in this table were repeated without the squared experience term, the linear teacher experience term was negative, and never had a t-statistic above 1.4 in absolute value, while the coefficient and level of significance of the homework variable was very little changed.

[7] At the mean level of homework assigned, which is 3.15 hours per week, all of the quadratic models suggest a positive marginal product of homework on test scores.

following year. As a test of robustness, models #1-3 were re-estimated on a subsample of observations where annual gains in test scores are less than or equal to +50% and greater than or equal to -33.3%. Thus if a student scored 60, 39, 60 in the first three years of testing the regression would exclude the first two observations from the regression, on the grounds that the test score of 39 was spuriously low. This restriction reduced the sample size by 4.4%. The results using this trimmed sample appear in columns #4-6. The coefficient on homework in the linear specification is about one quarter lower than in #2, but the overall fit of the model is much better. Although homework assigned remains by far the most significant predictor of test scores, teachers' experience and educational background become more significant in the trimmed sample. Thus there is some reason to believe that the estimated impact of homework hours on test scores may be overstated in the full-sample models due to the presence of outliers, and that the trimmed sample may give more reliable estimates.[8]

The above specifications do not conform fully to the theoretical model developed in Section 2. The model predicts that above a certain level of homework, a student will find it optimal to reduce his or her effort level, and hence learning, to the minimum possible. Thus homework should positively affect the student's test score up to some limit, and then have no effect after that point. Students who are better prepared at the start of the school year will find it easier to meet the homework standard imposed by the

[8] As an alternative to using a trimmed sample to reduce the effect of unusual changes in test scores, the full sample regressions were re-estimated by the method of least absolute deviations. This technique will reduce the impact of large outliers in the data. As shown in Table A-2 in the appendix, one finds coefficients and levels of significance that are much more like those found by OLS in the trimmed sample than the full sample. It appears that the trimmed data do effectively reduce the impact of unusual swings in test scores on the results.

teacher, so that the level of homework above which students "give up" should be a positive function of the students' raw ability and prior academic achievement. In order to measure the student's level of preparation at the start of the school year, I use the student's test score from the fall of the given year divided by the average test score for all students in that grade.

The hypothesis that there exists a threshold beyond which additional homework has no effect can then be tested by respecifying (4-1) as follows:

(4-2) $S_{ig} = \alpha + \beta S_{i,g-1} + \Gamma F_{ig} + \theta R_{ig} + \delta I(S_{i,g-1}, HW_{ig})HW_{ig} + \varepsilon_{ig}$

where I is a binary variable defined by

(4-3) $$I(S_{i,g-1}, HW_{ig}) = \begin{cases} 1 \text{ if } HW_{ig} < \gamma\left(\dfrac{S_{i,g-1}}{\bar{S}_{g-1}}\right), \text{ where } \gamma > 0 \\ 0 \text{ otherwise} \end{cases}$$

In this specification, the student's initial achievement is measured by his or her test score relative to the mean for all students in the previous grade, $S_{i,g-1} / \bar{S}_{g-1}$. If the amount of homework assigned is above the critical level given this initial achievement, then the impact of the homework on the student's achievement becomes zero. This is a model which is non-linear in the parameters. But it can be estimated by using a grid search over the parameter γ.

Table 2 presents the results of this grid search using both the full and trimmed samples. The models are identical to those used in Table 1 except for the respecification of the homework variable(s). The grid search was performed over values of γ ranging from 0, such that no students respond to homework, up to the value such that I=1 for all students in the sample, indicating that all students respond to homework. Two specifications of the impact of homework were used: linear and quadratic. In both the linear and quadratic specifications, the minimum of the sum of squared residuals was obtained for the same range of γ values. These results can be compared with those of Table 1, which are restricted versions of the non-linear model. A comparison of the results shows very little difference between the two tables, regardless of whether the full or trimmed samples are used. The hypothesis that the homework squared term can be excluded from the model is again retained. In the linear homework models, the coefficient and t-statistic on the homework variable are slightly higher than in the constrained models in Table 1. This difference is exactly what one would expect if a threshold does exist.

The reason for the similarity in results is that the data suggest that very few students are being given so much math homework that they 'give up' and exert no effort. As shown at the bottom of Table 2, in the full-sample models only one observation was above the threshold. For this observation, the student received substantially more than 3.150 hours of math homework per week, which was the average for the sample. The initial test scores for this observation was also only 44% of the average for students in the given grade.

One danger with the non-linear model is that its categorization of students into those who respond to homework and those who do not is sensitive to outliers in test scores. As a test of robustness, columns #3 and #4 of Table 2 show the non-linear model when re-estimated on the trimmed subsample. In these regressions, 0.77%, or 43 observations, are estimated to be beyond the threshold such that homework no longer induces effort. (These observations correspond to 43 different students distributed across all grade levels.) The overall fit of the model improves considerably. As in the full sample regression, the students who have "given up" have initial test scores below average for their grade, and levels of homework assigned which is more than twice the full-sample average of 3.15 hours per week.

Thus the most interesting conclusion from this analysis is that virtually the entire sample -- 99.3% -- exhibited a positive response to level of homework assigned. The finding that just under 1% of students do not respond positively to additional homework is a very optimistic finding. In a review of the literature on grade retention, Shepard and Smith (1989) report that approximately 2.6% to 8.9% of American students from kindergarten through Grade 12 fail a grade in any given year. Thus homework appears to affect effort positively even for a portion of the students who ultimately fail a grade.

More generally, it is important to know if the amount of homework assigned has a stronger impact on test scores of the better students within each grade. Several models were estimated to assess this possibility. The first model repeated the specification in Table 1 #2 but added a new term -- the interaction between hours of homework assigned and the student's relative test score $S_{i,g-1} / \overline{S}_{g-1}$. On both the full and trimmed samples, this interaction term was insignificant. Similarly, when Table 1 #2 was repeated with two additional terms, a dummy Q_1 indicating whether the student's initial test score was in the bottom quartile for the grade, and Q_1 interacted with homework, the interaction term was not statistically significant in either the full or trimmed sample. These exercises suggest that the impact of homework on students' rate of learning does not vary strongly with the student's standing relative to his or her peers across the country. The threshold model suggests a broadly similar picture, with only the bottom 1% of students failing to respond to homework.

Does the Effectiveness of Homework Vary with Respect to Other Student Traits?

The above finding suggests that the slope of the test score:homework relation does not vary with the student's initial achievement. But the marginal effectiveness of homework might vary with other personal traits. Using both the full and trimmed samples, and the specifications in Table 1 #2 and #5, the homework variable was entered on its own and was interacted with the following variables one at a time: a variable indicating year of study, the aforementioned measure of socioeconomic status of the family, and a dummy variable indicating whether both parents had 12 or fewer years of education (LOWED). In no case did any of these interactions become significant at even 10%. The most significant interacted regressor was LOWED, which in the trimmed sample was positive and had a t-statistic of 1.52, suggesting that homework might be slightly more effective with students whose parents are less well educated.

Robustness to Unobserved Heterogeneity

Omitted variable bias is a key concern in this model. Suppose that in general achievement of a student depends on the regressors already included in (4-1) as well as some unobserved trait of the student, ν_i:

(4-4) $S_{ig} = \alpha + \nu_i + \beta S_{i,g-1} + \Gamma F_{ig} + \theta R_{ig} + \delta HW_{ig} + \varepsilon_{ig}$

For example, if "ability" is multi-dimensional, then it is unlikely that the lagged test score will fully capture the ability of the student. It is conceivable that the omitted ability, ν_i, will be positively associated with the amount of homework which the teacher assigns, especially if the student's true ability is observable by the teacher. In this case, the coefficient on homework will be biased upward when (4-1) is estimated without the ν_i term.

It is not straightforward to estimate this model by treating ν_i as either a fixed effect or an error component, because the presence of the lagged dependent variable will in general cause inconsistent estimates when the number of observations per person is small (see e.g. Hsiao 1986).

An alternative is to respecify (4-4) so that the dependent variable becomes the change in test score from one year to the next:

(4-5) $S_{ig} - S_{i,g-1} = \alpha + \nu_i + \Gamma F_{ig} + \theta R_{ig} + \delta HW_{ig} + \varepsilon_{ig}$

and then to estimate the model by fixed effects. Such a model will be consistent even though the number of observations per person is small.

There is a second reason why it may be desirable to estimate model (4-5). If test scores are measured with error, then the inclusion of the lagged test score on the right hand side will bias *all* of the coefficients in the model. By instead specifying the dependent variable as the change in test score per year, model (4-5) removes this potential source of bias, but at the cost of higher variance in the error term, since the dependent variable is now in changes rather than in levels.[9] Thus the estimation of a model of test

[9] See Rogosa and Willett (1985) for a discussion of this and other reasons for why it may be preferable to estimate a growth model of learning.

score growth rather than the traditional value-added model is likely to lower precision of the estimates, but may remove inconsistency.

In economic terms, the fixed effect model (4-5) is useful because it ensures that the coefficient on homework will be identified only by *changes* in this variable for each student across years. If better students routinely get more homework, the estimates of (4-5) will be unbiased, whereas the earlier results, based on the value-added specification (4-1), is likely to be biased upward unless ability is perfectly observed. To the best of my knowledge, very few if any papers in the test score literature have been able to control for unobserved differences between students in this way.

Table 3 shows results from estimating (4-5). In general, the level of significance of the classroom traits falls, as is to be expected given that the dependent variable is now the change in test scores. But in both the full and trimmed samples, the hours of homework assigned remains highly significant. As will be shown later, the predicted cumulative effect of additional homework from these models is very close to the predicted gains from the more traditional form of the value-added models (4-1).[10]

As a check of robustness, it would be interesting to add a fixed effect to the original value-added model (4-1). But doing this is considerably more difficult owing to

[10] Another notable finding from the fixed effect models is that class size becomes negative and significant, unlike in the value-added models. One interpretation of this finding is that in general students who learn more quickly are assigned to larger classes, which biases the OLS coefficient on class size upward. Adding a fixed effect for each student may remove this bias. However, note that the effect of class size is quite small relative to that of homework. The predicted impact of an extra two hours of math homework per week is equivalent to the effect of reducing class size by approximately 6.8 to 14 students in the full and trimmed samples. These reductions in class size are quite sizable relative to the mean class size in the sample, of 24.8 students.

the presence of a lagged dependent variable. One approach to estimating a dynamic fixed effect model is to take the first difference of (4-4) to remove the fixed effect:

(4-6) $\Delta S_{ig} = \beta\Delta S_{i,g-1} + \Gamma\Delta F_{ig} + \theta\Delta R_{ig} + \delta\Delta HW_{ig} + \varepsilon_{ig} - \varepsilon_{ig-1}$

Due to correlation with the first differenced error term, the difference in lagged test score must be instrumented. As an instrument I use $S_{i,g-2}$, as suggested by Anderson and Hsiao (1981). This approach was attempted with the trimmed sample. Unfortunately, since three consecutive test scores are required per observation in this formulation, the maximum number of observations per student drops from two to one for students in the older cohort who were in Grade 10 in the first year, and from four to three in the younger cohort. Furthermore, if one test score is missing it can lead to up to three observations being dropped from the regression sample. Overall, the sample size dropped by just over half in the first-difference model. Consequently, the t-statistics fell on all of the coefficients. But homework assigned remained significant, with a coefficient of 0.180 and a t-statistic of 2.71. [11]

Although this latter method has the disadvantage of greatly reducing the sample size, it yields a similar conclusion to that obtained with the fixed effect model of test score growth in Table 3. Even after controlling for unobserved individual heterogeneity, additional math homework significantly improves students' rate of learning.

---

[11] Re-estimation of the first-difference IV model using the full sample gave a much lower coefficient estimate and t-statistic (0.11 and 1.34 respectively). But this result is hardly surprising, nor is it meaningful: the first differencing will aggravate measurement error due to the inclusion of a small number of outlier test score changes in this non-trimmed sample. To be more precise, the addition of 201 observations in which unusual changes in test scores occurred (growth above 50% or declines exceeding 33%) appears to be responsible for the lowered significance in this larger sample.

Assessment of the Relative Effectiveness of Extra Homework

The results in Tables 1 and 2 imply that an increase of one hour of homework per week increases test scores from one year to the next by about 0.3 to 0.4 point. The fixed effect models in Table 3 predict gains of 0.23 to 0.24 point. The least absolute deviation estimate in Table A-2 suggests a slightly lower effect of about 0.25 point. Thus a range of 0.23 to 0.4 seems reasonable.

The size of this effect is non-negligible. Consider the effect of increasing weekly math homework by three and one half hours, or one half hour per night. By the end of the school year the student's test score is predicted to rise from between 0.8 to 1.4 points depending on whether the coefficient of 0.23 or 0.4 is used. An improvement of 1.2 points is enough to bring a 45th percentile student up to the level of the 50th percentile student within a given grade.

The effect of homework becomes larger when one considers the cumulative effects of extra homework over time. For each of the regressions in Tables 1 and 2, the cumulative effect of an extra half an hour of math homework per night during Grades 7 to 11 was calculated, using the rate of depreciation of prior achievement implied by the coefficient on the lagged test score. The mean estimated net gain was 4.7 points.[12] Calculations based on the growth models in Table 3 yielded highly similar predicted gains of 4.0 and 4.2 points for the trimmed and full samples respectively. This predicted cumulative gain is very large relative to the average gain in test scores made by a given individual during the school year. In the full sample, the average year-to-year gain in math

---

[12] The lowest estimate was 4.1 points, obtained from Table 1, #5, which is the linear model without a threshold using the trimmed sample. The highest estimate was 5.1, which was the prediction from Table 1 #3 and Table 2 #2.

test scores was 2.2 points; for the trimmed sample it was 2.3. Thus giving students an extra half hour of homework per night during Grades 7 to 11 is estimated to raise their test scores by 1.7 to 2.1 grade equivalents. To put it another way, <u>a student who received no homework at all would have to spend approximately two extra years in school before learning as much as an identical student who had received half an hour of homework a night during Grades 7 through 11</u>.

These predicted gains are impressive, but are they credible? One way to answer this is to consider the extent to which half an hour of homework per night will increase the length of time which the student spends studying math. The LSAY asked teachers for the length of math classes for students for certain years. The average was 4.1 hours. For the regression sample, the mean amount of math homework assigned was 3.15, to yield a total of 7.25 hours of time on math per week on average. Thus 3.5 hours of extra homework per week would lengthen the time spent studying math by 48%. As a rough estimate, if we assumed that time spent on homework and in class were equally productive, and that this study time accounted for all of the 2.2 to 2.3 gain in test scores per year on average, then a 48% increase in time spent studying math should increase annual gains by about 1.1 points. Remarkably, the predicted annual gains from this extra homework, as calculated above using the estimates from Tables 1 through 3, fall in the range 0.8 to 1.4. Thus the predicted gains in test scores from extra homework appear to be entirely reasonable.

The above exercise implies that time spent on homework may be as effective as time spent in the classroom. To test this intriguing hypothesis, models #3 and #5 from Table 1 were repeated with an additional regressor: the teacher's report of the number of hours of math class as an additional regressor. The results appear in Table A-3. Relative to the results in Table 1, the coefficient on hours of homework rises somewhat and remains highly significant. The coefficient on hours of class time is very close to that for homework, but slightly larger, and is also highly significant. Perhaps most interesting, a test for the equality of the coefficients of hours of homework and hours of class time retains the null, as shown by the p-values at the bottom of the table. Remarkably, time spent on homework and time spent in the classroom appear to be equally productive. [13] [14]

The predicted gains from additional homework are also impressive relative to the measured effectiveness of other school inputs. In contrast to the 4.7 point gain from homework, reducing class size by one during these grades leads to only a 0.07 point gain on average based on Tables 1 and 2, and even these meager estimates depend on using the lower bound of the 95% confidence intervals in these tables. The more optimistic fixed effect results in Table 3 suggest that reducing class size by 1 results in a cumulative gain in test scores of 0.15 to 0.35 point. Similarly, the cumulative gain from using teachers with a Master's degree is only 1.6 points, and the effect is significant only in some specifications. (The fixed effect models suggest that teachers with Master's degrees are mildly *less* effective.)

---

[13] Unfortunately, the question about length of math classes was available for only 61% of the samples used in Table 1, mainly because the question was not asked in the first year of the survey. For this reason, hours of math class was not included as a regressor in the other tables in this paper. The relatively small number of observations probably also explains why the significance of some of the other school variables drops in Table A-3 relative to Table 1.

[14] Model (4-5), the test score fixed effect specification, was also re-estimated with hours of math class time added. In this model neither class time nor hours of homework were significant. But this is understandable given the small sample size once one adds hours of class time, combined with the fact that the addition of a fixed effect for each person annihilates all of the information for the many people in the restricted sample for whom there was only one observation. (Fully 61.6% of the people in this subsample had only one observation, and so the information on these people did not contribute to the identification of the effects of homework.)

Not only are the cumulative effects of improving class size and teacher education relatively small, but implementation of such changes is likely to be fairly expensive. For instance, a standard log wage regression for teachers in the March 1993 Current Population Survey suggests that school teachers who hold Master's degrees command roughly 17% more than their counterparts without a Master's degree. Similarly, assuming optimistically that the only cost of decreasing class size is added instructional expenditures, a reduction in the national teacher-pupil ratio from its 1988 value of 17.3 to 16.3 would have increased the cost of public primary and secondary education by $6.2 billion in that year.[15]

Would Additional Homework Increase Costs by Necessitating Smaller Classes?

In contrast with reductions in class size and the hiring of more highly educated teachers, assigning more homework appears to have a greater impact without added expense. But is additional homework truly a policy prescription which entails no additional school spending? In particular, homework assignments may trigger learning only if teachers collect and grade each assignment. If such were the case, then it may not be feasible to increase the number of homework assignments without reducing class size to prevent teachers from becoming overworked.

The LSAY allows a direct answer to this question, since each teacher answers the question "What percentage of the homework assignments did you correct and return to students?". Table 4 accordingly repeats the original specification (4-1) in column #1, and

[15] On the more conservative assumption that decreasing class size increases the need for other inputs such as building stock and support staff proportionately, the cost balloons to $10.6 billion. These calculations are based on National Center for Education Statistics, (1991), pages 70 and 153.

then in column #2 replaces the hours of homework *assigned* with the hours of homework which are *assigned, corrected and returned*. The latter is the product of hours of work assigned and the proportion graded. If homework is effective only if teachers grade it, then model 2 should fit better than model 1. In fact, model 2 does not fit the data as well: it appears that homework assigned is a more significant predictor of student achievement than the hours of homework assigned and graded.

Models 1 and 2 are non-nested, and thus cannot be tested against each other using traditional methods. However, a useful specification test of both models is to perform a J-test to test whether the fitted values from one model explain any of the variation in test scores that is not already explained by the other model. (See Davidson and MacKinnon, 1981.) This non-nested test is *not* a model selection test since it is possible that both models could be rejected.

In the simple case considered here, where both models are linear and there is only one additional regressor in the two candidate models, the J test can be performed by adding the hours of homework graded to the original model, and testing that this variable does not belong in the original model. The results from estimating this artificially nested model on the trimmed sample appear in column #3 of Table 4. The specification test of the original model, in column #1, retains the null. (The p-value, shown in the bottom line of the table, is simply the p-value on the t-statistic for hours of homework graded.) However, the non-nested test of model #2, in which homework hours is replaced by hours of graded homework, rejects this model.

Columns #4-6 in Table 4 repeat the analysis using the fixed effect model of test score growth (4-5). The results are highly similar, and stronger in that the p-value for the test that the model with hours of homework assigned is correct is 0.934, indicating that this model can explain virtually all of the variation explained by the model with the hours of homework which are graded.

The implication of these non-nested tests is that it is hours of homework *assigned*, and not hours of homework graded, which influences student achievement. Thus it need not be the case that in order to increase homework loads policymakers would have to reduce class size substantially.[16]

One may wonder how it could be that extra homework appears to increase student effort even if not all of the homework is collected and graded. One answer may be that teachers rely on uncertainty in order to induce student effort, by not stating to the class in advance which homework assignments will be graded. Another plausible explanation is that parents monitor whether their children are completing their homework assignments. Indirect evidence supporting this idea comes from the fact that in the regression sample in Table 1, 79.2% of students reported that 'My parents insist I do my homework'.

The optimistic conclusion that extra homework assignments do not require smaller classes gains support from the following less formal exercise. If teachers who have larger classes assign less homework, it would imply that teachers already face severe time constraints. Accordingly, extra homework assignments would probably necessitate

[16] Identical tests using the full sample came to highly similar conclusions. The p-values for the test of the original model and the model with the amount of homework graded and returned were 0.1273 and 0.0002 respectively, based on the models in #1 and #2. For the growth models in columns #4 and #5, the p-values were 0.871 and 0.034.

reductions in class size, which would make for a very expensive policy reform. But the evidence points to the contrary conclusion. The correlation between class size on the one hand, and homework assigned and the hours of homework which are corrected and returned on the other hand, is *positive*: 0.096 and 0.062 respectively. Thus, teachers with larger classes actually assign and grade slightly *more* hours of homework than do teachers in smaller classes.[17] The implication is clear: it need not be the case that a national policy of heavier homework loads must be accommodated by decreasing class size in order to prevent teachers from facing heavier workloads. At the least, teachers with smaller classes should be able to increase homework assignments to the levels given by teachers with larger classes. Furthermore, Table 4 has shown that it is sufficient to assign more homework without grading all of it. Thus a policy of heavier homework need not increase teachers' work hours excessively.[18]

Estimating the Impact of Math Homework on Wages

The LSAY does not obtain information on workers' wages after leaving school. But an estimate of the impact of homework on wages can be made using estimates of the impact of math scores on wages from other data-sets. Recent work by Grogger and Eide (1995) and Murnane, Willett and Levy (1995), using High School and Beyond (HSB) and the National Longitudinal Study of the High School Class of 1972 (NLS72), find that the

[17] Crude regressions of hours of homework assigned on dummies for grade level and class size suggest a small effect: increasing class size by 10 is predicted to increase weekly hours of homework by 20 minutes.
[18] The correlation coefficients reported in this paragraph are calculated after demeaning the variables by grade level. Using raw data, similar correlations result: 0.085 and 0.056 respectively.

impact of math scores on wages six years after graduation rose substantially between the 1970's and 1980's.

Using these studies to calculate the impact of homework on wages can yield at best a rough approximation, because the math tests in HSB and the NLS72 on the one hand and the LSAY on the other are not identical.[19] The mean test score among Grade 12 students who took the math test in HSB in 1980 was 12.69 out of 25, while the mean test score of Grade 12 students in the LSAY was 65.40 out of 100, or in rescaled terms, 16.35 out of 25. Thus a conservative assumption might be that an increase of 1 point in the LSAY test (rescaled to a total of 25) is equivalent to a gain of 12.69/16.35=0.776 point in the HSB test.[20] Murnane, Willett and Levy (1995) report that in their analysis HSB the mean coefficient on the math score in a regression of log hourly wages in 1986 was 0.014.[21] Thus the 4.7 point predicted gain in LSAY math scores from an extra half an hour of math homework per night was converted into a predicted gain in log hourly wages as 4.7*0.25*0.776*0.014=0.0128, or a gain in wages of about 1.3%. Using the more conservative predicted gain of 4.0 points in LSAY math test scores, an extra half hour of math homework nightly throughout Grades 7 to 11 is predicted to yield wage gains of

---

[19] The HSB math test used by Murnane, Willett and Levy consisted of 25 questions which asked students to choose the higher of two quantities. The LSAY math test consisted of roughly 50-60 questions taken from the National Assessment of Educational Progress which were designed to test students' ability to measure, perform basic numerical operations, and do algebra and geometry. No higher mathematical skills such as calculus were tested. Murnane, Willett and Levy (1995) report that the HSB math questions focused on arithmetic, following directions and interpreting graphs, but did not involve algebra or geometry. Both the HSB and LSAY test scores were scaled using the logistic model employed in Item Response Theory to ensure comparability across tests. Thus the LSAY and HSB tests are not identical but bear a strong resemblance.

[20] An alternative interpretation is that student performance rose between 1980, when HSB started, and the late 1980's when the LSAY was conducted. But I opt for the former interpretation because it will lead to lower, more conservative, estimates of the impact of math homework on wages.

[21] This figure is the simple mean of the estimated coefficients for men and women.

1.1%. These are meaningful gains. Moreover, they are likely to understate the true gains because the log hourly wage regressions of Murnane, Willett and Levy condition on educational attainment. It seems reasonable that students with higher test scores will obtain more education and thus increase their earnings through this second indirect route.

5. Conclusions

This work makes several contributions to the literature. First, it extends the existing theoretical paradigm, in which student achievement is modeled as depending on years of schooling and various measures of school resources, to one in which the effort expected of students also affects their performance. Second, in an empirical application of the model, the paper finds that homework standards have a strong positive effect on students' test scores. Extra homework appears to have a larger and more significant impact on student achievement than the more standard spending-related measures of school characteristics such as teachers' education and experience or class size. Remarkably, time spent on homework appears to be as effective as time spent in the classroom. Third, the non-linear model allows an estimate of what percentage of the student population in the country is receiving 'too much' homework. This percentage is found to be positive but less than 1%, suggesting that math teachers could increase almost all students' achievement by assigning more homework than they do at present.

Perhaps the most convincing estimates in the paper model test score growth in a fixed effect framework. This specification controls for unobserved traits of each student, and ensures that the homework coefficient will be identified solely by changes in the amount of homework assigned across classes for each student. The fixed effect model yields similar conclusions to the earlier results: an extra half hour of math homework per night in Grades 7 to 11 is estimated to advance a student almost two full grade equivalents. Rough estimates suggest that this could translate into a wage gain of at least 1.1-1.3%.

Student achievement in math appears to respond to hours of homework assigned more strongly than to the hours of homework which are assigned and subsequently collected and graded by the teacher. Thus, a policy of increasing homework standards need not overburden teachers. These findings promise a low-cost method for improving school quality.

International comparisons of public schools suggest that much could be done to increase the amount of homework which American students currently receive. In a comparison of public schools in the United States, Japan, China and Taiwan, Stevenson and Stigler (1992) find that American students performed much more poorly than did their foreign counterparts on standardized tests, and at the same time spent less time on homework each day.

More broadly speaking, the paper suggests that much of the earlier literature on the education production function has focused too narrowly on the role of financial resources in determining student achievement, while ignoring non-financial inputs into education. In particular, we may be able to learn much by explicitly modeling how the choices of students respond to the standards which schools and teachers impose. It appears that in public education, as in so many other spheres of human activity, incentives do matter.

Figure 1
Educational Achievement by Ability

The dotted line indicates the distribution of educational achievement after an increase in the educational standard.

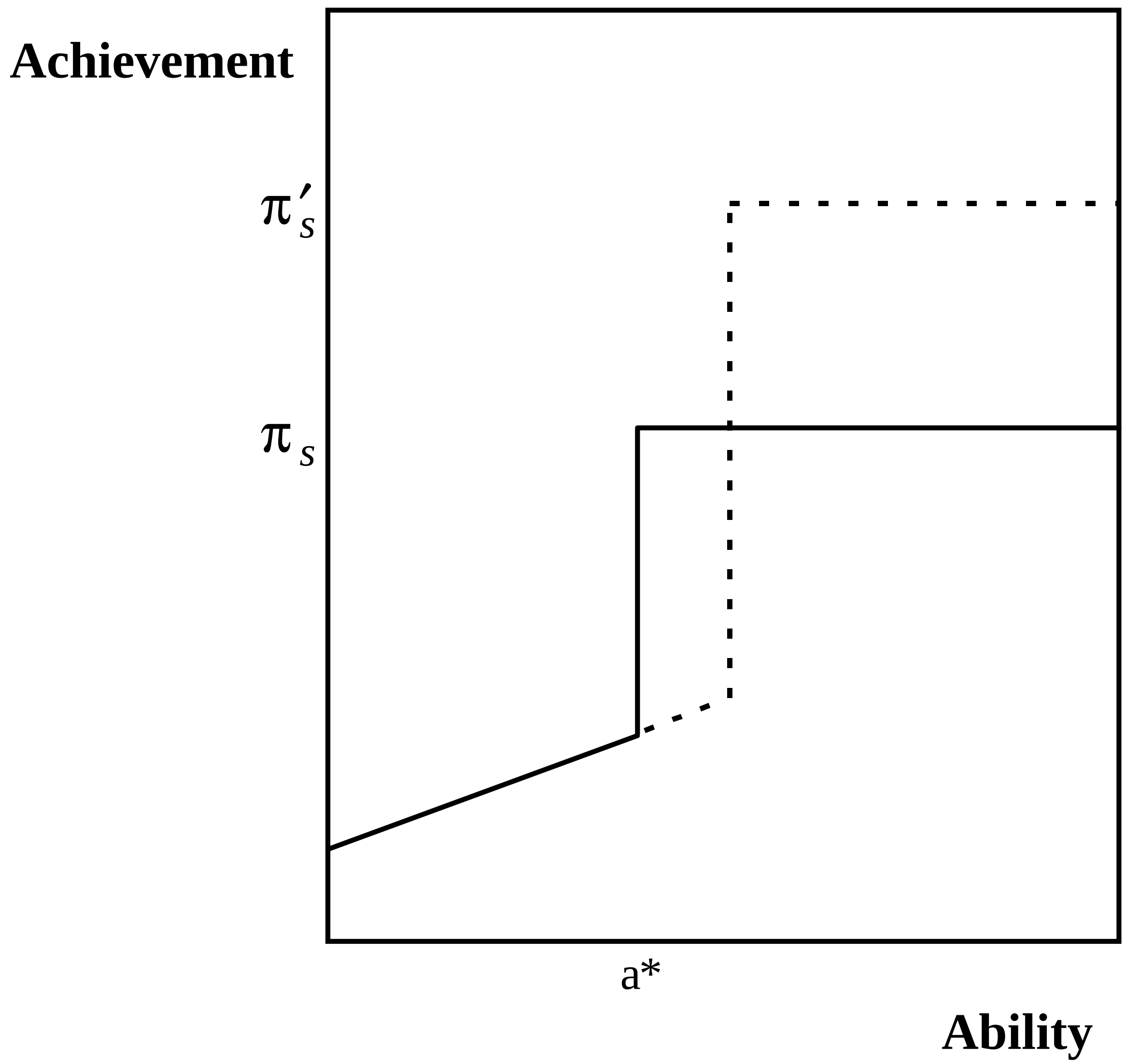

Table 1

Ordinary Least Squares Estimators of Value-Added Math Test Score Equations

Other regressors not shown are listed in the text. The row at the bottom provides probability values for the stated F test. The trimmed regressions use a sample which deletes an observation if the observed gain in test scores in the given year is greater than 50% or less than -33.3%. This reduces the sample by 4.4%. (Sample size is 6352 in regressions 1-3 and 6073 in regressions 4-6.) T-statistics are White-corrected for heteroskedasticity.

Regression #	**1**	**2**	**3**	**4**	**5**	**6**
Regression Sample		**Full**			**Trimmed**	
Lagged Score	0.7763	0.7654	0.7651	0.8747	0.8659	0.8655
	(65.17)	(62.47)	(62.42)	(99.43)	(96.40)	(96.06)
Full-Time Exp.	-0.1008	-0.0917	-0.0941	-0.1108	-0.1037	-0.1069
	(-2.18)	(-1.99)	(-2.03)	(-2.91)	(-2.73)	(-2.79)
F.-T. Exp. Squared	0.0026	0.0024	0.0024	0.0031	0.0029	0.0030
	(1.82)	(1.65)	(1.69)	(2.69)	(2.53)	(2.59)
Associate Degree	0.2923	0.3475	0.3464	0.5035	0.5446	0.5431
	(0.86)	(1.03)	(1.03)	(1.78)	(1.93)	(1.92)
Master's Degree	0.5422	0.4143	0.4260	0.5347	0.4377	0.4540
	(2.03)	(1.55)	(1.59)	(2.43)	(1.99)	(2.05)
Ph.D.	0.3897	0.2369	0.1974	-1.5357	-1.6507	-1.7045
	(0.19)	(0.12)	(0.10)	(-0.80)	(-0.87)	(-0.89)
Class Size	0.0099	0.0050	0.0046	0.0088	0.0054	0.0048
	(0.68)	(0.35)	(0.32)	(0.77)	(0.47)	(0.42)
School Enrollment	-0.0009	-0.0008	-0.0008	-0.0009	-0.0009	-0.0009
	(-2.81)	(-2.71)	(-2.71)	(-3.79)	(-3.73)	(-3.74)
% Full Lunch Assist.	0.0128	0.0136	0.0134	-0.0002	0.0001	-0.0002
	(1.11)	(1.17)	(1.15)	(-0.02)	(0.01)	(-0.02)
Hours of Homework (HW)		0.4156	0.5040		0.3075	0.4310
		(5.54)	(2.46)		(5.13)	(2.53)
HW Squared			-0.0108			-0.0151
			(-0.46)			(-0.79)
R-Squared	0.6013	0.6033	0.6033	0.7254	0.7266	0.7266
Adjusted R-Squared	0.5991	0.6011	0.6010	0.7238	0.7250	0.7250
F Test HW, HW Sq.			0.00000			0.00000

Table 2

Non-Linear Models with a Threshold for the Effect of Homework on Math Test Scores

Other regressors not shown are listed in the text. The bottom of the table lists the percentage of observations for which students are estimated to be at a corner solution, and the mean test scores and levels of homework for these students. The trimmed regressions are as described in Table 1. (Sample size is 6352 in regressions 1 and 2 and 6073 in regressions 3 and 4.) T-statistics are White-corrected for heteroskedasticity.

Regression #	**1**	**2**	**3**	**4**
Sample	**Full**	**Full**	**Trimmed**	**Trimmed**
Lagged Score	0.7651	0.7649	0.8624	0.8625
	(62.36)	(62.33)	(94.79)	(94.55)
Full-Time Exp.	-0.0917	-0.0938	-0.1139	-0.1131
	(-1.99)	(-2.02)	(-3.01)	(-2.97)
F.-T. Exp. Squared	0.0024	0.0024	0.0032	0.0032
	(1.65)	(1.69)	(2.76)	(2.72)
Associate Degree	0.3474	0.3464	0.5287	0.5276
	(1.03)	(1.03)	(1.87)	(1.87)
Master's Degree	0.4152	0.4257	0.4918	0.4874
	(1.55)	(1.59)	(2.24)	(2.22)
Ph.D.	0.2364	0.2011	-1.7548	-1.7359
	(0.12)	(0.10)	(-0.92)	(-0.91)
Class Size	0.0049	0.0046	0.0052	0.0055
	(0.34)	(0.31)	(0.45)	(0.48)
School Enrollment	-0.0008	-0.0008	-0.0009	-0.0009
	(-2.72)	(-2.72)	(-3.66)	(-3.66)
% Full Lunch Assist.	0.0135	0.0133	-0.0008	-0.0007
	(1.16)	(1.14)	(-0.08)	(-0.07)
Hours of Homework (HW)	0.4189	0.4982	0.3549	0.3006
	(5.56)	(2.42)	(5.72)	(1.70)
HW Squared		-0.0097		0.0071
		(-0.41)		(0.34)
R-Squared	0.6033	0.6034	0.7269	0.7269
Adjusted R-Squared	0.6011	0.6010	0.7252	0.7252
F Test HW, HW*HW		0.00000		0.00000
γ	17.25-22.85	17.25-22.85	9.27-9.33	9.27-9.33
% of Obs. at Corner Solution	0.02%	0.02%	0.71%	0.71%
Mean HW, Corner-Solution Observations	10.00	10.00	8.14	8.14
Mean $S_{i,g-1} / \overline{S}_{g-1}$, Corner-Solution Obs.	0.44	0.44	0.77	0.77

Table 3

Models of Test Score Gains with Individual Fixed Effects

The dependent variable is the one year change in math test scores. Specifications include the same regressors as in previous models, except for the deletion of the lagged test score. T-statistics are White-corrected for heteroskedasticity. See Table 1 for sample sizes.

Regression #	**1**	**2**
Regression Sample	**Full**	**Trimmed**
Full-Time Exp.	-0.0572	-0.0751
	(-0.87)	(-1.42)
F.-T. Exp. Squared	0.0012	0.0022
	(0.59)	(1.39)
Associate Degree	0.0936	0.0864
	(0.20)	(0.23)
Master's Degree	-0.6466	-0.2847
	(-1.73)	(-0.96)
Ph.D.	-0.9807	-3.6151
	(-0.41)	(-1.76)
Class Size	-0.0699	-0.0326
	(-3.38)	(-2.14)
School Enrollment	-0.0009	-0.0006
	(-1.25)	(-1.07)
% Full Lunch Assist.	-0.0375	-0.0592
	(-0.99)	(-1.90)
Hours of Homework (HW)	0.2384	0.2284
	(2.30)	(2.70)
R-Squared	0.4262	0.4703
Adjusted R-Squared	-0.2193	-0.1523

Table 4

Comparison of Specifications with Hours of Homework Assigned and Hours of Homework Which are Graded, and Non-Nested Tests

Specifications are OLS models with the same regressors as in previous models. Results for the trimmed sample are presented in the table. Sample size is 5892. T-statistics are White-corrected for heteroskedasticity. Models 3-6 include fixed effects for each student.

Regression #	**1**	**2**	**3**	**4**	**5**	**6**
Dependent Variable		**Score**			**Change in Score**	
Lagged Score	0.8641	0.8707	0.8641			
	(94.42)	(96.69)	(94.37)			
Full-Time Exp.	-0.1101	-0.1209	-0.1099	-0.0797	-0.0944	-0.0796
	(-2.82)	(-3.10)	(-2.81)	(-1.44)	(-1.69)	(-1.42)
F.-T. Exp. Squared	0.0031	0.0033	0.0031	0.0023	0.0026	0.0023
	(2.59)	(2.75)	(2.59)	(1.34)	(1.50)	(1.33)
Associate Degree	0.5470	0.4690	0.5493	0.2231	0.1984	0.2237
	(1.90)	(1.62)	(1.90)	(0.56)	(0.50)	(0.56)
Master's Degree	0.3781	0.4656	0.3775	-0.4555	-0.3427	-0.4557
	(1.69)	(2.09)	(1.68)	(-1.48)	(-1.12)	(-1.48)
Ph.D.	-1.6485	-1.7089	-1.6422	-3.6084	-3.4185	-3.6077
	(-0.86)	(-0.90)	(-0.86)	(-1.74)	(-1.66)	(-1.74)
Class Size	0.0016	0.0050	0.0015	-0.0290	-0.0263	-0.0290
	(0.13)	(0.42)	(0.13)	(-1.80)	(-1.63)	(-1.80)
School Enrollment	-0.0009	-0.0008	-0.0009	-0.0001	-0.0001	-0.0001
	(-3.48)	(-3.36)	(-3.47)	(-0.22)	(-0.23)	(-0.22)
% Full Lunch Assist.	0.0011	-0.0004	0.0012	-0.0494	-0.0507	-0.0494
	(0.11)	(-0.04)	(0.12)	(-1.40)	(-1.44)	(-1.40)
Hours of Homework (HW)	0.3170		0.3208	0.2504		0.2513
	(5.12)		(4.41)	(2.78)		(2.44)
Hours of Graded HW		0.1470	-0.0064		0.1104	-0.0015
		(2.66)	(-0.10)		(1.41)	(-0.02)
R-Squared	0.7242	0.7234	0.7242	0.4807	0.4800	0.4807
Adjusted R-Squared	0.7225	0.7217	0.7225	-0.1637	-0.1653	-0.1641
Non-Nested J Test p-value	0.884	0.00003		0.934	0.014	

Table A-1
Means and Standard Deviations of Variables

Sample corresponds to that used in Table 1, regressions 1-3.

Variable	Mean	σ
Test Score	60.109	13.309
Lagged Test Score	57.894	12.360
Grade 7	0.213	0.410
Grade 8	0.187	0.390
Grade 9	0.176	0.381
Grade 10	0.300	0.458
Grade 11	0.124	0.330
Father Education < 12 Years	0.124	0.329
Fath. Ed. High School Diploma	0.496	0.500
Fath. Ed. Some College	0.107	0.309
Fath. Ed. 4-Year Degree	0.157	0.363
Fath. Ed. Advanced Degree	0.117	0.321
Mother Education < 12 Years	0.114	0.318
Moth. Ed. High School Diploma	0.611	0.488
Moth. Ed. Some College	0.089	0.284
Moth. Ed. 4-Year Degree	0.139	0.346
Moth. Ed. Advanced Degree	0.047	0.211
Both parents Ed. <= 12 years	0.527	0.499
Male	0.494	0.500
Suburban	0.425	0.494
Rural	0.366	0.482
Urban	2.158	0.742
Black	0.097	0.296
Hispanic	0.079	0.270
Asian	0.022	0.147
Native American	0.013	0.113
North-East	0.184	0.388
North-Central	0.362	0.481
South	0.324	0.468
West	0.130	0.336
Socio-economic Status	0.054	0.770
Score/(Class Ave. Score)	1.032	0.202
Teacher Experience	14.409	8.165
Teacher Exp. Squared	275.723	252.737
Teacher: Assoc. Degree	0.143	0.346
Teacher: Master's Degree	0.539	0.494
Teacher: Ph.D. Degree	0.003	0.054
Math Class Size	24.834	7.613

Continued on next page

Table A-1, Continued

Variable	Mean	σ
Enrollment in School	883.131	488.541
% Receiving Full Lunch Assistance	14.771	14.960
% Black	0.127	0.202
% Hispanic	0.027	0.065
% Asian	0.022	0.056
Hours of Math Class Time (for sample in Table A-3 #1)	4.098	0.744
Hours of Math Homework per Week	3.150	1.556
Hours of Math Homework Graded and Returned per Week	2.058	1.689

Table A-2

Least Absolute Deviation Estimates Based on Full Sample

The regressors are the same as in specifications #1-3 in Table 1. Number of observations is 6352.

Regression #	**1**	**2**	**3**
Lagged Score	0.8543	0.8443	0.8406
	(112.44)	(109.28)	(109.89)
Full-Time Exp.	-0.0902	-0.0849	-0.0753
	(-3.53)	(-3.31)	(-2.89)
F.-T. Exp. Squared	0.0021	0.0021	0.0018
	(2.16)	(2.12)	(1.79)
Associate Degree	0.3182	0.2848	0.3302
	(1.40)	(1.25)	(1.45)
Master's Degree	0.4529	0.2924	0.3250
	(2.54)	(1.64)	(1.80)
Ph.D.	-1.2004	-1.5941	-1.7010
	(-0.86)	(-1.14)	(-1.21)
Class Size	0.0031	0.0025	-0.0001
	(0.31)	(0.25)	(-0.01)
School Enrollment	-0.0009	-0.0008	-0.0008
	(-4.02)	(-3.73)	(-3.53)
% Full Lunch Assist.	-0.0030	-0.0008	-0.0031
	(-0.40)	(-0.11)	(-0.37)
Hours of Homework (HW)		0.2540	0.4959
		(4.89)	(4.13)
HW Squared			-0.0284
			(-1.69)
R-Squared	0.5990	0.6010	0.6010
Adjusted R-Squared	0.5968	0.5987	0.5986

Table A-3

Replication of the Basic Model in Table 1 with Hours of Math Classroom Time as an Additional Regressor

The p-value reported at the bottom of the table is for the test that the coefficients on Hours of Math Homework and Hours of Classroom Time are equal. T-statistics are White-corrected for heteroskedasticity.

Regression Sample	**Full**	**Trimmed**
Lagged Score	0.7839	0.8908
	(51.67)	(81.15)
Full-Time Exp.	-0.0269	-0.0538
	(-0.45)	(-1.09)
F.-T. Exp. Squared	0.0010	0.0023
	(0.54)	(1.52)
Associate Degree	0.8075	0.7817
	(1.68)	(1.96)
Master's Degree	0.2603	0.1693
	(0.76)	(0.60)
Ph.D.	0.3609	-1.6907
	(0.17)	(-0.84)
Class Size	-0.0084	-0.0001
	(-0.48)	(-0.01)
School Enrollment	-0.0008	-0.0008
	(-2.06)	(-2.41)
% Full Lunch Assist.	0.0350	0.0143
	(2.19)	(1.07)
Hours of Homework (HW)	0.4925	0.3498
	(5.56)	(4.91)
Hours of Class Time	0.5771	0.3742
	(3.04)	(2.29)
R-Squared	0.6028	0.7315
Adjusted R-Squared	0.5990	0.7288
Number of Obs.	3892	3710
P-Value	0.696	0.863

References

Anderson, T. W., and Hsiao, Cheng. "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association* 76 (September 1981): 598-606.

Becker, William, and Rosen, Sherwin. "The Learning Effect of Assessment and Evaluation in High School." Discussion Paper 90-7, Economics Research Center, NORC, 1990.

Betts, Julian R. "Which Types of Public School Spending are Most Effective? New Evidence on the School Quality Debate." Discussion Paper 95-03, Dept. of Economics, University of California, San Diego, 1995.

---. "Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature." In *Does Money Matter? The Link between Schools, Student Achievement, and Adult Success*, edited by Gary Burtless. Washington, D.C.: Brookings Institution, forthcoming.

Card, David, and Alan B. Krueger, “The Economic Returns to School Quality: A Partial Survey,” Working Paper #334, Industrial Relations Section, Princeton University, 1994.

Coleman, James et al. *Equality of Educational Opportunity*. Washington: Government Printing Office, 1966.

Cooper, Harris. *Homework*. New York, NY: Longman, 1989.

Costrell, Robert M. "A Simple Model of Educational Standards." *American Economic Review* 84 (September 1994): 956-71.

Davidson, Russell, and MacKinnon, James G. "Several Tests for Model Specification in the Presence of Alternative Hypotheses." *Econometrica* 49 (May 1981): 781-93.

Grogger, Jeff, and Eide, Eric. "Changes in College Skills and the Rise in the College Wage Premium." *Journal of Human Resources* 30 (Spring 1995): 280-310.

Hanushek, Eric A. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24 (September 1986): 1141-77.

Hsiao, C. *Analysis of Panel Data*. Cambridge: Cambridge University Press, 1986.

Kang, Suk. "A Formal Model of School Reward Systems." In *Incentives, Learning, and Employability*, edited by John H. Bishop. Columbus, Ohio: National Center for Research in Vocational Education, Ohio State University, 1985.

Keith, Timothy Z. et al. "Parental Involvement, Homework, and TV Time: Direct and Indirect Effects on High School Achievement." *Journal of Educational Psychology* 78 (October 1986): 373-80.

Miller, J. D. et al. *LSAY Codebook: Student, Parent, and Teacher Data for Cohort One for Longitudinal Years One through Four (1987-1991)*. Vol. 1. DeKalb, IL: Northern Illinois University, 1992.

Murnane, Richard J.; Willett, John B.; and Levy, Frank. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77 (May 1995): 251-66.

National Center for Education Statistics. *Digest of Education Statistics 1991*. Washington, D.C.: U.S. Government Printing Office, 1991.

Rogosa, David R., and Willett, John B. "Understanding Correlates of Change by Modeling Individual Differences in Growth." *Psychometrika* 50 (1985): 203-28.

Shepard, Lorrie A. and Smith, Mary Lee. “Introduction and Overview.” In *Flunking Grades: Research and Policies on Retention*, edited by Lorrie A. Shepard and Mary Lee Smith. New York: The Falmer Press, 1989.

Stevenson, Harold W., and Stigler, James W. *The Learning Gap: Why our Schools are Failing and What We Can Learn from Japanese and Chinese Education*. New York: Summit Books, 1992.

Walberg, Herbert J.; Fraser, Barry J.; and Welch, Wayne W. "A Test of a Model of Educational Productivity among Senior High School Students." *Journal of Educational Research* 79 (January/February 1986): 133-39.