The Impact of Homework on Student Achivement

Ozkan Eren*
Department of Economics
Southern Methodist University

Daniel J. Henderson†
Department of Economics
State University of New York at Binghamton

April 27, 2006

Abstract

Utilizing parametric and nonparametric techniques, we asses the role of a heretofore relatively unexplored "input" in the educational process, homework, on academic achievement. Our results indicate that homework is an important determinant of student test scores. Relative to more standard spending related measures, extra homework has a larger and more significant impact on test scores. However, the effects are not uniform across different subpopulations; we find additional homework to be most effective for high and low achievers. Moreover, the parametric estimates of the educational production function overstate the impact of schooling related inputs. In all estimates, the homework coefficient from the parametric model maps to the upper deciles of the nonparametric distribution and as a by-product the parametric model understates the percentage of students with negative responses to additional homework.

Keywords: Generalized Kernel Estimation, Nonparametric, School Inputs, Stochastic Dominance, Student Achivement
JEL Classification: C14, I21, I28

*Ozkan Eren, Department of Economics, Southern Methodist University, Dallas, TX 75275-0496. Tel: (214) 768-3957. Fax: (214) 768-1821. E-mail: oeren@smu.edu.

†Corresponding author: Daniel J. Henderson, Department of Economics, State University of New York, Binghamton, NY 13902-6000, Tel: (607) 777-4480, Fax: (607) 777-2681, E-mail: djhender@binghamton.edu.

1 Introduction

The stagnation of student achievement over the past three decades in the United States is well documented (for example, see Epple and Romano 1998 and Hoxby 1999), despite the fact that real spending per pupil has risen at a remarkably steady pace of 3.5% per year over the period 1890-1990 (Hanushek 2003) and that aggregate public expenditure on primary and secondary education amounts to approximately $200 billion (Betts 2001). Given the discontinuity that exists between educational expenditures and student achievement, economists have produced a voluminous body of research attempting to explore the primary influences of student learning. The vast majority of papers in this area have focused on spending related "inputs" such as class size and teachers' credentials. With a few exceptions, these studies conclude that measured school "inputs" have only limited effects on student outcomes (Hanushek 2003). In light of these pessimistic findings, it is surprising how little work has been devoted to understanding the impact of other aspects of the educational environment on student achievement.[1] In particular, given parental concerns, policy debates and media interest (for example, see Ratnesar 1999), very little empirical research to date has been completed on the role of homework in student achievement (theoretical papers include Betts 1998, Costrell 1994, and Neilson 2005).

We know of two empirical studies that examine the effects of homework on student outcomes. Aksoy and Link (2000), using the National Educational Longitudinal Study of 1988 (NELS:88) data, find positive and significant effects of homework on tenth grade math test scores. However, the authors rely on student responses on the hours of homework, which carries the potential risk of a spurious correlation because it may reflect unobserved variation in student ability and motivation. Betts (1998) presents the only empirical work that, to our knowledge, focuses on the hours of homework assigned by the teacher. This measure of homework is actually a policy variable, which the school or the teacher can control. Using the Longitudinal Study of American Youth data and panel estimations, Betts obtains a substantial effect of homework on math test scores. Specifically, an extra half hour of math homework per night in grades 7 to 11 is estimated to advance a student nearly two grade equiv-

[1]Notable exceptions are Betts and Grogger (2003) and Figlio and Lucas (2003), who analyze the impact of grading standards on student achievement.

alents. Furthermore, in a nonlinear model setting, the author argues that virtually all students (99.3% of the sample) could benefit from extra homework and thus math teachers could increase almost all students achievement by assigning more homework.

Although the aforementioned papers provide careful and important evidence on the effects of homework, there are numerous gaps remaining. First, there may be heterogeneity in the returns to homework. Theoretical treatments of the topic indicate that the responses to extra homework will depend on a student's ability level (for example, see Betts 1997 and Neilson 2005). In this respect, the impact of homework may differ among students of different achievement levels. Second, the existing educational production function literature relies mostly on parametric regression models. Although popular, parametric models require several stringent assumptions. In particular, the errors are generally assumed to come from a specified distribution and the functional form of the educational production function is given *a priori.* Given that the theory predicts a non-monotonic relation between homework and student achievement, a parametric specification which fully captures the true relation may be difficult to find. Further, if the functional form or distributional assumptions do not hold, the parametric model will most likely lead to biased estimates.

Given these potential shortcomings, we adopt a nonparametric approach. Nonparametric estimation procedures relax the functional form assumptions associated with the traditional parametric regression model and create a tighter fitting regression curve through the data.[2] These procedures do not require assumptions on the distribution of the error nor do they require specific assumptions on the form of the underlying production function (for example, see Polachek, Kniesner and Harwood 1978). Furthermore, the procedures generate unique coefficient estimates for each observation for each variable. This attribute enables us to estimate the return to homework for each student and make inference regarding heterogeneity in the returns.

Utilizing the NELS:88 data, we reach four striking empirical findings. First, relative to more standard spending related measures such as class size, extra homework appears to have a larger and more significant impact on mathematics achievement.

[2]Nonparametric estimation has been used in other labor economics domains to avoid restrictive functional form assumptions, for example, see Cobb-Clark and Hildebrand (2004), Henderson, Olbrecht, and Polachek (2006) and Kniesner and Li (2002).

However, the effects are not homogenous among different subpopulations. We find additional homework to be most effective for high and low achievers. Second, including the teacher's evaluation of the overall class achievement is crucial in the estimations. In the absence of such a control, the schooling inputs are upward biased. Third, time spent on homework and time spent in class do not seem to be equally productive. This may suggest that learning by doing is a more effective tool for improvement in student achievement. Finally, the parametric estimates of the educational production function overstate the impact of schooling related inputs. In particular, both the homework and class size coefficients from the parametric model map to the upper deciles of the distribution of nonparametric estimates. Moreover, the parametric model understates the percentage of students with negative responses to an additional hour of homework.

The remainder of the paper is organized as follows. Section 2 describes the nonparametric estimation strategy, as well as the statistical tests used in the paper. The third section discusses the data while the fourth presents the results. Finally, Section 5 concludes.

2 Empirical Methodology

2.1 Parametric Model

To begin, we estimate a parametric specification of the educational production function

$$TS_{ikmj} = f(HW_m, W_i, C_k, T_m, \xi_j, \beta) + \varepsilon_{ikmj}, \tag{1}$$

where TS is the test score of student i in school j in class k and HW denotes the hours of homework assigned by teacher m. The vector W represents individual and family background characteristics, as well as ex ante achievement (lagged test scores), C is a vector of class inputs and T is a vector of teacher characteristics. We control for all factors invariant within a given school with the fixed effect ξ, β is a vector of parameters to be estimated and ε is a zero mean, normally distributed error term. Our main parameter of interest is the coefficient on homework, which represents the effect of an additional hour of homework on student test scores.

2.2 Generalized Kernel Estimation

Parametric regression models require one to specify the functional form of the underlying Data Generating Process (DGP) prior to estimation. Correctly specified parametric models provide consistent estimates and inference based on such estimates is valid. However, uncertainty exists about the functional form of the educational production function because the theory does not provide a guide as to an appropriate functional form (for example, see Betts 1997, Hanushek 2003, and Todd and Wolpin 2003). There could be nonlinear relations as well as interactions among regressors, which standard parametric models may not capture.

Given the potential shortcomings of the parametric model, we also estimate a nonparametric version of (1). To proceed, we utilize Li-Racine Generalized Kernel Estimation (Li and Racine 2004 and Racine and Li 2004) and express the test score equation as

$$TS_i = m(x_i) + \eta_i, \qquad i = 1, ..., N \tag{2}$$

where $m(\cdot)$ is the unknown smooth educational production function, η_i is an additive error term and N is the sample size. The covariates of equation (1) are subsumed in $x_i = [x_i^c, x_i^u, x_i^o]$, where x_i^c is a vector of continuous regressors (for example, hours of homework), x_i^u is a vector of regressors that assume unordered discrete values (for example, race), x_i^o is a vector of regressors that assume ordered discrete values (for example, parental education).

Taking a first-order Taylor expansion of (2) with respect to x_j yields

$$TS_i \approx m(x_j) + (x_i^c - x_j^c)\beta(x_j) + \eta_i, \tag{3}$$

where $\beta(x_j)$ is defined as the partial derivative of $m(x_j)$ with respect to x^c. The estimator of $\delta(x_j) \equiv \binom{m(x_j)}{\beta(x_j)}$ is given by

$$\begin{aligned} \widehat{\delta}(x_j) &= \begin{pmatrix} \widehat{m}(x_j) \\ \widehat{\beta}(x_j) \end{pmatrix} = \left[\sum_{i=1}^{N} K_{\widehat{h}} \begin{pmatrix} 1 & (x_i^c - x_j^c) \\ (x_i^c - x_j^c) & (x_i^c - x_j^c)\,(x_i^c - x_j^c)' \end{pmatrix} \right]^{-1} \\ & \sum_{i=1}^{N} K_{\widehat{h}} \begin{pmatrix} 1 \\ (x_i^c - x_j^c) \end{pmatrix} TS_i, \end{aligned} \tag{4}$$

where $K_{\widehat{h}} = \prod_{s=1}^{q} \left(\widehat{\lambda}_s^c\right)^{-1} l^c\left(\frac{x_{si}^c - x_{sj}^c}{\widehat{\lambda}_s^c}\right) \prod_{s=1}^{r} l^u\left(x_{si}^u, x_{sj}^u, \widehat{\lambda}_s^u\right) \prod_{s=1}^{p} l^o\left(x_{si}^o, x_{sj}^o, \widehat{\lambda}_s^o\right)$. K_h is the commonly used product kernel (Pagan and Ullah 1999), where l^c is the standard normal kernel function with window width $\lambda_s^c = \lambda_s^c(N)$ associated with the s^{th} component of x^c. l^u is a variation of Aitchison and Aitken's (1976) kernel function for unordered categorical data and l^o is the Wang and Van Ryzin (1981) kernel function for ordered categorical data.

Estimation of the bandwidths $h = (\lambda^c, \lambda^u, \lambda^o)$ is typically the most salient factor when performing nonparametric estimation. For example, choosing a very small bandwidth means that there may not be enough points for smoothing and thus we may get an undersmoothed estimate (low bias, high variance). On the other hand, choosing a very large bandwidth, we may include too many points and thus get an oversmoothed estimate (high bias, low variance). This trade-off is a well known dilemma in applied nonparametric econometrics and thus we resort to automatic determination procedures to estimate the bandwidths. Although there exist many selection methods, one popular procedure (and the one used in this paper) is that of Least-Squares Cross-Validation. In short, the procedure chooses $(\lambda^c, \lambda^u, \lambda^o)$ which minimize the least-squares cross-validation function given by

$$CV(\lambda^c, \lambda^u, \lambda^o) = \frac{1}{N}\sum_{j=1}^{N}[TS_j - \widehat{m}_{-j}(x_j)]^2, \tag{5}$$

where $\widehat{m}_{-j}(\cdot)$ is the commonly used leave-one-out estimator of $m(x)$.[3]

2.3 Model Selection Criteria

To assess the correct estimation strategy, we utilize the Hsiao, Li, and Racine (2003) specification test for mixed categorical and continuous data. The null hypothesis is that the parametric model $(f(x_i, \beta))$ is correctly specified $(H_0 : \Pr[\mathrm{E}(TS_i|x_i) = f(x_i, \beta)] = 1)$ against the alternative that it is not $(H_1 : \Pr[\mathrm{E}(TS_i|x_i) = f(x_i, \beta)] < 1)$. The test statistic is based on $I_N \equiv \mathrm{E}\left(\mathrm{E}(\varepsilon|x)^2 f(x)\right)$, where $\varepsilon = y - f(x, \beta)$. I_N is non-negative and equals zero if and only if the null is true. The resulting test statistic is

$$J_N = \frac{N\left(\widehat{\lambda}^c\right)^{q/2}\widehat{I}_N}{\widehat{\sigma}_N} \sim N(0,1), \tag{6}$$

[3] All bandwidths in this paper were calculated using N ©.

where

$$\widehat{I}_N = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \widehat{\varepsilon}_i \widehat{\varepsilon}_j K_{\widehat{h}},$$

$$\widehat{\sigma}_N^2 = \frac{2\left(\widehat{\lambda}^c\right)^q}{N^2} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \widehat{\varepsilon}_i^2 \widehat{\varepsilon}_j^2 K_{\widehat{h}}^2,$$

$K_{\widehat{h}}$ is the product kernel, $\widehat{\lambda}^c$ $(\widehat{h})$ is the optimally chosen bandwidth for the continuous (complete set of) covariates, and q is the number of continuous regressors. If the null is false, J_N diverges to positive infinity. Unfortunately, the asymptotic normal approximation performs poorly in finite samples and a bootstrap method is generally suggested for approximating the finite sample null distribution of the test statistic. This is the approach we take.

2.4 Stochastic Dominance

Nonparametric estimation as described in equation (4) generates unique coefficient estimates for each observation for each variable. This feature of nonparametric estimation enables us to compare (rank) the returns to homework for subgroups and make inferences about who benefits most from an additional hour of homework. Empirical examination of such comparisons based on stochastic dominance (SD) is the approach utilized in this paper.[4] The comparison of the effectiveness of a policy on different subpopulations based on a particular index (such as conditional mean) is highly subjective; different indices may yield substantially different conclusions. In contrast, finding a SD relation provides uniform ranking regarding the impact of the policy among different groups and offers robust inference through the distribution.

To proceed, let W and V denote two outcome variables being compared (W and V might refer to the coefficients on homework obtained from the nonparametric estimates for males and females, respectively). $\{w\}_{i=1}^{N_0}$ is a vector of N_0 observations for W; $\{v\}_{i=1}^{N_1}$ is an analogous vector of realizations of V and $F(w)$ and $G(v)$ represent the cumulative distribution functions (CDF) of W and V, respectively.

Consider the null hypotheses of interest as

[4] For an empirical application of SD in the school quality literature, see Maasoumi, Millimet, and Rangaprasad (2005).

Equality of Distributions :

$$F(z) = G(z) \quad \forall z \in \Omega. \tag{7a}$$

First Order Stochastic Dominance : F dominates G if

$$F(z) \leq G(z) \quad \forall z \in \Omega, \tag{7b}$$

where Ω is the union support for W and V. To test the null hypotheses, we define the empirical CDF for W as

$$\hat{F}(w) = \frac{1}{N_0} \sum_{i=1}^{N_0} I(W \leq w), \tag{8}$$

where I denotes the indicator function and $\hat{G}(v)$ is defined similarly. Next, we define the following Kolmogorov-Smirnov statistics

$$T_{eq} = \sup_{z \in \Omega} |\hat{F}(z) - \hat{G}(z)| \tag{9a}$$

$$T_{fsd} = \sup_{z \in \Omega} (\hat{F}(z) - \hat{G}(z)) \tag{9b}$$

for testing the equality and first order stochastic dominance (FSD) relation, respectively.

Unfortunately, the asymptotic distributions of these nonparametric sample based statistics under the null are generally unknown because they depend on the underlying distributions of the data. We need to approximate the empirical distributions of these test statistics to overcome this problem. Our strategy following Abadie (2002) is as follows:

(i) Let T be a generic notation for T_{eq} and for T_{fsd}. Compute the test statistics T for the original sample of $\{w_1, w_2.....w_{N_0}\}$ and $\{v_1, v_2.....v_{N_1}\}$.

(ii) Define the pooled sample as $\Omega = \{w_1, w_2.....w_{N_0}, v_1, ...v_{N_1}\}$. Resample $N_0 + N_1$ observations with replacement from Ω and call it $\hat{\Omega}$. Divide $\hat{\Omega}$ into two groups to obtain $\hat{T}$.

(iii) Repeat step (ii) B times.

(iv) Calculate the p-values of the tests with p-value $= \sum_{b=1}^{B} I(\hat{T}_b > T)/B$. Reject the null hypotheses if the p-value is smaller than some significance level α, where $\alpha \in (0, 1/2)$.

By resampling from Ω, we approximate the distribution of the test statistics when $F(z) = G(z)$. Note that for (7b), $F(z) = G(z)$ represents the least favorable case for the null hypothesis. This strategy allows us to estimate the supremum of the probability of rejection under the composite null hypothesis, which is the conventional definition of test size.[5]

3 Data

The data is obtained from the National Educational Longitudinal Study of 1988 (NELS:88), a large longitudinal study of eighth grade students conducted by the National Center for Educational Statistics (NECS). The NELS:88 is a stratified sample, which was chosen in two stages. In the first stage, a total of 1032 schools on the basis of school size were selected from a universe of approximately 40,000 schools. In the second stage, up to 26 students were selected from each of the sample schools based on race and gender. The original sample contains approximately 25,000 eighth grade students. Follow-up surveys were administered in 1990, 1992, 1994 and 2000.

To measure academic achievement, students were administered cognitive tests in reading, social sciences, mathematics and science during the base year (eighth grade), first follow-up (tenth grade) and second follow-up (twelfth grade). Each of the four grade specific tests contain material appropriate for each grade, but included sufficient overlap from previous grades to permit measurement of academic growth. Although four test scores are available per student, teacher and class information sets (discussed below) are only available for two subjects per student.

We utilize tenth grade math test scores as our dependent variable in light of the findings of Grogger and Eide (1995) and Murnane, Willett, and Levy (1995).[6] These studies find a substantial impact of mathematics achievement on postsecondary education, as well as on earnings. Our variable of interest is the *hours of homework*

[5]Ideally one would reestimate the nonparametric returns within each bootstrap replication to take into account the uncertainty of the returns. Unfortunately, this would require reestimating the bandwidths for each bootstrap replication, which would be extremely computationally difficult, if not impossible. Thus, the bootstrapped p-values most likely differ slightly from their 'true' values. Nonetheless, if we obtain a large p-value, it is unlikely that accounting for such uncertainty would alter the inference.

[6]We follow Boozer and Rouse (2001) and Altonji, Elder, and Taber (2003) and utilize item response theory math test scores.

assigned daily and comes directly from the student's math teacher reports. As noted in Betts (1997) and Neilson (2005), this measure of homework is a policy variable, which the school administrator or the teacher can control. Relying on hours spent on homework from the student reports is not as accurate and may yield spurious correlations because it may reflect unobserved variation in student ability and motivation.

Given that researchers interested in the impact of school quality measures are typically (and correctly) concerned about the potential endogeneity of school quality variables, we utilize a relatively lengthy vector of student, family, teacher and classroom characteristics. The NELS:88 data enables us to tie teacher and class-level information directly to individual students and thus circumvents the risk of measurement error and aggregation bias. Furthermore, we include school fixed effects as described in equations (1) and (2) to capture differences between schools that may affect student achievement. Specifically, our estimations control for the following variables:

Individual: gender, race, lagged (eighth grade) math test score;

Family: father's education, mother's education, family size, socioeconomic status of the family;

Teacher: gender, race, age, education;

School: school fixed effects;

Class: class size, number of hours the math class meets weekly, teacher's evaluation of the overall class achievement.

Information on individual and family characteristics are obtained from the base year survey questionnaires and data pertaining to the math teacher and class comes from the first follow-up survey. Observations with missing values for any of the variables defined above are dropped. We further restrict the sample to students who attend public schools. Table 1 reports the weighted summary statistics of some of the key variables for the 6913 students in the public school math sample and for the regression sample used for estimation.[7] The means and standard deviations in the

[7] Our regressions do not use weights. Instead we include controls for the variables used in the stratification (see Rose and Betts 2004 for a similar approach).

regression sample are strikingly similar to those obtained when using the full set of potential public school observations. This similarity provides some assurance that missing values have not distorted our sample.

4 Empirical Results

4.1 Parametric Results

Our parametric specifications are presented in Table 2. For all regression estimates, White standard errors are reported beneath each estimate. The first column of Table 2 gives a large significant coefficient for homework. An additional hour of homework is associated with a gain of 4.20 (0.58) points in math achievement. Given that the mean test score is approximately 51.56, this represents an increase of slightly above eight percent. However, this model is simplistic in that it does not take into account many observable variables that are known to affect test scores. In the second column of Table 2, we include demographic and family characteristics. There is a slight decrease in the homework coefficient. The third column adds student's eighth grade math scores, which provide an important control for ex ante achievement and captures all previous inputs in the educational production process, giving the results a "value-added" interpretation (for example, see Goldhaber and Brewer 1997, Hanushek 1979, and Todd and Wolpin 2003). Including the student's 8th grade math score greatly reduces the homework coefficient from 3.56 (0.50) to 0.97 (0.20). However, the coefficient is still statistically significant.

An important concern regarding the effect of homework and any other school quality variables is that schools may differ in both observable and unobservable dimensions. If school traits are correlated with homework or other inputs, then it is likely that the coefficients will be biased. Therefore, it is most prudent to control for any observed and unobserved factors common to all students in a school. We do so by including the school fixed effects in the fourth column of Table 2. The school dummies are jointly significant (p-value = 0.00), but the homework coefficient remains practically unchanged.

The fifth and sixth columns of Table 2 add teacher and classroom characteristics (class size and weekly hours of math class), respectively. Even though the effect of homework is similar in magnitude, two points are noteworthy regarding the selected covariate estimates. First, the class size coefficient is positive and statistically significant, in that, increasing the number of students in a math class from the sample average of 23 to 33 will lead to an increase of 0.45 points in math scores. This finding is consistent with Goldhaber and Brewer (1997), who use the NELS:88 data to assess the impact of class size on tenth grade math test scores. Second, in contrast to Betts (1997), we do not find a significant effect of weekly hours of math class on test scores. Moreover, the coefficient is substantially small in magnitude. The null hypothesis of equality between the coefficients on the hours of homework and hours of class time is strongly rejected. This indicates that time spent on homework and time spent in class are not equally productive (Aksoy and Link 2000 reach a similar conclusion). It appears that time spent on homework is what matters.

As noted above, the school fixed effects will capture any factors common to all students in a school, but there may still be some unobserved ability differences across students within a school. For instance, if the overall ability of students in a class is high due to nonrandomness in the assignments of students to classes, then the teacher may increase (decrease) the homework load for students in that particular class. If this is the case, the homework coefficient is going to be upward (downward) biased. To control for this possibility, we utilize the teacher's questionnaire on the overall achievement level of the math class, which is split into four categories: high, average, low, and widely differing. Regression estimates controlling for class achievement are given in the seventh column of Table 2. The class achievement variables are jointly significant (p-value = 0.00). The homework coefficient is still statistically significant, but considerably diminished in magnitude. A similar reduction is observed in the class size effect as well and is now only marginally significant.

Finally, in the last column of Table 2, we tested the potential nonlinear effects of homework in the parametric specification by adding a quadratic term. In this model, the homework squared term is negative and statistically significant, suggesting evidence for diminishing returns to the amount of homework assigned. The return to homework becomes zero at around 3.15 hours per day and is negative afterwards. This corresponds to 0.3% of the sample. At the mean level of hours of homework, which is 0.64 per day, the marginal product of homework is roughly 0.97.[8]

Overall, our parametric specifications provide four key insights. First, inclusion of the teacher's evaluation of class achievement in the regression estimates is crucial. In the absence of such a control, the coefficients on homework and class size are overstated. Further, we believe that teacher's assessment of class ability purges out ability differences within a school to an extent and alleviates any bias arising from the possible endogeneity of homework. This is important given the difficulty finding a valid instrument with this particular data set. Besides, Goldhaber and Brewer (1997) find that unobservable school, teacher and class effects do not appear to be correlated with observed school quality variables and thus do not cause biased estimates of these variables. Second, time spent on homework and time spent in class do not seem to be equally productive. This may suggest that learning by doing is a more effective tool for improvement in student achievement. Third, compared to more standard spending related measures such as class size, additional homework appears to have a larger and more significant impact on math test scores. Fourth, hours of homework assigned exhibit diminishing returns, but only 0.3% of the sample respond negatively to additional homework.

4.2 Nonparametric Results

Prior to discussing the results, we conduct the Hsiao, Li, and Racine (2003) specification test based on the assumption that the correct functional form is the last column of Table 2. The parametric model is strongly rejected (p-value = 0.00). The linear parametric model (seventh column of Table 2) is also strongly rejected (p-value =

[8] In addition to our parametric specifications presented in Table 2, we estimate the last column by: (i) Including two additional variables regarding the teacher's treatment of homework. The NELS:88 asked teachers whether they keep records of who turned in the assignments and whether they return the assignments with grades or corrections; each split into four categories: all the time, most of the time, some of the time and never. The homework coefficients are qualitatively identical. (ii) Including the average 8th grade math score at the school level as an additional control variable to capture pre-high school peer effects. The homework coefficients are virtually identical and the average 8th grade math score at the school level appears to be negative and highly insignificant with a value equal to -0.01 (0.07). (iii) Including a host of school control variables, rather than utilizing the fixed effect estimation. The school specific variables are urban/rural status, region, total school enrollment, grade-level enrollment, student racial composition and the percentage of students receiving free lunch. The OLS estimates are 1.28 (0.36) and -0.21 (0.08) for homework and homework squared, respectively.

0.00). These finding raises concerns regarding the functional form assumptions of the educational production function in the existing school quality literature.

Turning to the results, Table 3 displays the nonparametric estimates (which control for individual, family, teacher, classroom characteristics and school fixed effects) of homework on math test scores. Given the number of parameters obtained from the Generalized Kernel Estimation procedure, it is tricky to present the results. Unfortunately, no widely accepted presentation format exists. Therefore, in Table 3, we give the mean estimate, as well as the coefficients at each decile of the distribution along with their respective bootstrapped standard errors. The mean nonparametric estimate is positive but statistically insignificant with a value of 0.57 (0.35). Looking at the coefficient distribution, we observe positive and significant effects for the upper four deciles. The explained portion of the variance of student achievement rises from 0.83 to 0.93 when we switch from the parametric to nonparametric model. Precision set aside, the parametric estimate at the mean level of homework obtained from the last column of Table 2 is larger than the corresponding mean of the nonparametric estimate. More importantly, more than 20% of the nonparametric estimates are negative, whereas those with negative responses to homework are only 0.3% of the sample from the parametric model. Table A1 in Appendix A displays the sample statistics for those with negative homework coefficients. The most interesting pattern, when we compare it with the regression sample, is observed in the overall class achievement. Students with negative coefficients are intensified in classes, which the teacher evaluates as average. We further analyze this point in the next section.

Table 4 presents the nonparametric estimates of the remaining continuous covariates of Table 2. We present the mean, as well as the nonparametric estimates corresponding to the 25^{th}, 50^{th} and 75^{th} percentiles of the coefficient distribution (labelled Q1, Q2 and Q3). The results for the eighth grade math scores are in line with the parametric estimates and are statistically significant throughout the distribution. The class size effect, however, differs from the parametric estimates. The mean nonparametric estimate indicates a reversal in the sign of the class size effect. Even though we do obtain primarily negative coefficients, a majority are insignificant and thus we are unable to draw a definite conclusion at this point. We had hoped to find an unambiguous conclusion on the effect of the class size given the conflicting results in the literature (see Hanushek 2003).

The mean return to time spent in class is negative and larger in magnitude than the parametric estimate. In addition, the negative effect is statistically significant at the first quartile. The mean nonparametric estimate for family size is similar in magnitude to the parametric model. However, the coefficient for the socioeconomic status differs substantially. The parametric and (mean) nonparametric estimates are 0.23 (0.18) and 0.86 (0.29), respectively (the full set of estimates for these last two variables are available upon request).

In sum, relaxation of the functional form assumptions in the educational production model reveal at least three findings. First, at the mean level, the predicted effect of homework from the parametric estimate (0.97) is roughly 1.7 times larger than the nonparametric estimate (0.57). Second, parametric estimates understate the percentage of students with negative responses to homework. However, extra homework continues to be effective for at least 40% of the sample under the nonparametric model. Third, the sign of the (mean) class size coefficient is reversed from positive to negative and is no longer significant.

4.3 Effects of Homework by Achievement Group

Given the concentration of students with negative responses at the average achievement level, we further explore the impact of homework on subgroups based on the teacher's evaluation of the class. Table 5 displays the mean nonparametric estimate, the coefficients at each decile of the distribution, as well as the parametric estimate of homework for each subgroup. In the parametric specifications, we exclude the homework squared term unless it is significant.[9]

The first column of Table 5 presents the results for the high achievement group. The parametric estimate of homework is significant with a value of 2.02 (0.56) and is higher than the corresponding 90^{th} percentile of the nonparametric estimates; the mean nonparametric counterpart is 0.77 (0.36) and is statistically significant. Thus, the nonparametric model indicates that the parametric model *vastly* overstates the homework effect for virtually the entire subsample. In addition, the parametric model

[9] In contrast to the parametric model, we do not need to split the sample and reestimate for each subgroup because we have already obtained a unique coefficient for homework for each individual in the nonparametric model.

cannot capture the heterogeneity inherent in the model. For instance, the homework effect is more than twice as large at the 90^{th} percentile (1.92) as it is at the 60^{th} percentile (0.86).

The second column presents the estimates for the average achievement group. The parametric *and* nonparametric estimates do not indicate any significant effect of homework on math test scores. Even though the coefficients are insignificant, the nonparametric model indicates that nearly 40% of the subsample responded negatively to extra homework. This may not be surprising given that the students with negative responses are intensified in average achievement classes.

For the low achievement group, unlike the first two columns, we include the statistically significant homework squared term in the parametric specification. The return to homework becomes zero at around 2.04 hours and is negative afterwards. This corresponds to roughly 0.35% of the subsample. At the mean level of homework, which is 0.52 hours per day, the marginal product of homework is roughly 1.69 and is higher than the corresponding 80^{th} percentile of the nonparametric estimates. The mean nonparametric estimate is 0.79 (0.43) and marginally significant. Similar to the first column, the parametric model overstates the homework effect and moreover, understates the percentage of students with negative responses, which is more than 12% of the subsample based on the nonparametric estimates.

For completeness, the last column presents the estimates for students in classes with widely differing ability levels. The coefficients are large in magnitude but are only statistically significant for the upper two deciles of the nonparametric estimates.

Table 6 displays the results for three continuous covariates for each subgroup. We present the parametric results, as well as the nonparametric mean estimates and the nonparametric estimates corresponding to the 25^{th}, 50^{th} and 75^{th} percentiles of the coefficient distribution. Three results emerge. First, the parametric estimates of eighth grade math scores are similar in magnitude to the mean (median) of the nonparametric estimates. Second, for three of the subgroups (high, average and widely differing), different from the parametric estimates, we observe predominantly negative coefficients for the class size effect, but the coefficients are statistically insignificant. For the low achievement group, however, the class size effect is positive and significant. The parametric estimate lies in the upper extreme tail of the corresponding distribution of nonparametric estimates. Specifically, the parametric estimate, 0.13 (0.05), maps to roughly the 90^{th} percentile of the nonparametric estimates. This difference between groups may explain why the literature finds such different conclusions regarding class size. It further stresses the importance of using a methodology which allows for heterogeneity in the returns. Finally, for the average achievement group, the nonparametric estimates of time spent in class are negative and statistically significant for the first quartile, mean, and median.

The final set of results are provided in Table 7. We report the p-values associated with the null hypotheses of equality and FSD for the homework coefficient distributions among the four subgroups. The corresponding CDFs are plotted in Figure 1. For all subgroups, we can easily reject equality of distributions at conventional confidence levels (p-value = 0.00). In terms of rankings, homework return for the three subgroups (high, low and widely differing), dominate average achievers' returns in the first order sense and further confirm that extra homework is less effective or may not be effective at all for average achievers. We do not observe FSD between the widely differing ability group and low or high achievers. There is some evidence of FSD for the return distribution of low achievers over high achievers, however, it is not statistically significant (p-value = 0.26).[10]

The theoretical models of homework (for example, see Betts 1997 and Neilson 2005) suggest that beyond a certain level of homework, a student will find it optimal to reduce his or her effort level to the minimum. Thus, homework should positively affect the student's achievement up to some limit and then have no effect. In this respect, extra homework leading gains for high achievers is not at odds with theory. The mean hours of homework for high achievers is 0.74 (0.40), but this amount may be far away from the subgroup's "give-up" limit. The potential puzzle in our results is that extra homework is not effective for average achievers, despite leading gains for low achievers. One possibility is that average achievers are at the edge of their maximum effort, whereas low achievers are below their threshold level. The mean hours of homework are 0.64 (0.38) and 0.52 (0.38) for average and low achievers, respectively. If the "give-up" level for low achievers is some value greater than 0.52,

[10] We also estimate the return to homework for subgroups based on gender and race. The coefficients are insignificant in most cases. The mean estimates, SD tests and CDFs are given in Appendix B.

then they will benefit from the extra homework. Although this is by no means a definitive explanation for our findings, it is a plausible explanation.

5 Conclusion

The stagnation of academic achievement in the United States and elsewhere has given rise to a growing literature seeking to understand the determinants of student learning. Utilizing parametric and nonparametric techniques, we assess the impact of a heretofore relatively unexplored "input" in the educational process, homework, on tenth grade test performance.

Our results indicate that homework is an important determinant in student achievement. Relative to more standard spending related measures such as class size, extra homework appears to have a larger and more significant impact on math test scores. However, the effects are not uniform across different subpopulations. We find additional homework to be most effective for high and low achievers. Next, time spent on homework and time spent in class do not seem to be equally productive. This may suggest that learning by doing is a more effective tool for improvement in student achievement. Finally, parametric estimates of the educational production function overstate the impact of schooling related inputs and thus raises concerns regarding the commonly used specifications in the existing literature. Specifically, the sign of the mean class size effect is reversed from positive to negative and is no longer significant when we switch to a nonparametric model. In all estimates, both homework and class size coefficients from the parametric model map to the upper deciles of the nonparametric coefficient distribution. Moreover, parametric estimates understate the percentage of students with negative responses to an additional hour of homework.

From a policy point of view, it is premature to conclude that extra homework will yield Pareto improvements in educational outcomes. On one hand, homework helps low achievers to catch up, but on the other hand, additional homework may increase the performance gap between the best and average students. Therefore, a better understanding of the complexity of student responses to homework is required.

References

[1] Abadie, A. (2002), "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284-292.

[2] Aitchison, J. and C.G.G. Aitken (1976), "Multivariate Binary Discrimination by Kernel Method," *Biometrika*, 63, 413-420.

[3] Aksoy, T. and C.R. Link (2000), "A Panel Analysis of Student Mathematics Achievement in the US in the 1990s: Does Increasing the Amount of Time in Learning Activities Affect Math Achievement?" *Economics of Education Review,* 19, 261-277.

[4] Altonji, J., T. Elder, and C. Taber (2003), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113, 151-184.

[5] Betts, J.R. (1997), "The Role of Homework in Improving School Quality," Unpublished Manuscript, Department of Economics, University of California, San Diego.

[6] ———. (1998), "The Impact of Educational Standards on the Level and Distribution of Earnings," *American Economic Review*, 88, 266-275.

[7] ———. (2001), "The Impact of School Resources on Women's Earnings and Educational Attainment: Findings from the National Longitudinal Survey of Young Women," *Journal of Labor Economics*, 19, 635-657.

[8] ——. and J. Grogger (2003), "The Impact of Grading Standards on Student Achievement, Educational Attainment and Entry-Level Earnings," *Economics of Education Review*, 22, 343-352.

[9] Boozer, M.A. and C. Rouse (2001), "Intraschool Variation in Class Size: Patterns and Implications," *Journal of Urban Economics*, 50, 163-189.

[10] Cobb-Clark, D.A. and V. Hildebrand (2004), "The Wealth of Mexican Americans," *Journal of Human Resources*, forthcoming.

[11] Costrell, R.M. (1994), "A Simple Model of Educational Standards," *American Economic Review*, 84, 956-971.

[12] Epple, D. and R. Romano (1998), "Competition Between Private and Public Schools, Vouchers, and Peer Group Effects," *American Economic Review*, 88, 33-62.

[13] Figlio, D.N. and M.E. Lucas (2003), "Do High Grading Standards Affect Student Performance?" *Journal of Public Economics*, 88, 1815-1834.

[14] Goldhaber, D.B. and J.B. Brewer (1997), "Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity," *Journal of Human Resource*, 32, 505-523.

[15] Grogger, J. and E. Eide (1995), "Changes in College Skills and the Rise in the College Wage Premium," *Journal of Human Resources*, 30, 280-310.

[16] Hanushek, E.A. (1979), "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *Journal of Human Resources*, 14, 351-388.

[17] ———. (2003), "The Failure of Input Based Schooling Policies," *Economic Journal*, 113, 64-98.

[18] Henderson, D.J., A. Olbrecht and S. Polachek (2006), "Do Former Athletes Earn More at Work? A Nonparametric Assessment," *Journal of Human Resources*, forthcoming.

[19] Hoxby, C.M. (1999), "The Productivity of Schools and Other Local Public Goods Producers," *Journal of Public Economics*, 74, 1-30.

[20] Hsiao, C., Q. Li and J. Racine (2003), "A Consistent Model Specification Test with Mixed Categorical and Continuous Data," *Journal of Econometrics*, forthcoming.

[21] Kniesner, T.J. and Q. Li (2002). "Nonlinearity in Dynamic Adjustment: Semiparametric Estimation of Panel Labor Supply," *Empirical Economics*, 27, 131-48.

[22] Li, Q. and J. Racine. (2004), "Cross-Validated Local Linear Nonparametric Regression," *Statistica Sinica*, 14, 485-512.

[23] Maasoumi, E., D.L. Millimet and V. Rangaprasad (2005), "Class Size and Educational Policy: Who Benefits from Smaller Classes?" *Econometric Reviews*, 24, 333-368.

[24] Murnane, R.J., J.B. Willett and F. Levy (1995), "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, 77, 251-266.

[25] N ©, Nonparametric software by Jeff Racine (http://www.economics.mcmaster.ca/racine/).

[26] Neilson, W. (2005), "Homework and Performance for Time-Constrained Students," *Economics Bulletin*, 9, 1-6.

[27] Pagan, A. and A. Ullah (1999), *Nonparametric Econometrics*, Cambridge, Cambridge University Press.

[28] Polachek, S.W., T.J. Kniesner and H.J. Harwood (1978), "Educational Production Functions," *Journal of Educational Statistics*, 3, 209-231.

[29] Racine, J. and Q. Li (2004), "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data," *Journal of Econometrics*, 119, 99-130.

[30] Ratnesar, R. (1999), "The Homework Ate my Family," *Time,* January 25, 54-63.

[31] Rose, H. and J.R. Betts (2004), "The Effect of High School Courses on Earnings," *Review of Economics and Statistics*, 86, 497-513.

[32] Todd, P.E. and K.I. Wolpin (2003), "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal*, 113, F3-F33.

[33] Wang, M.C. and J.V. Ryzin (1981), "A Class of Smooth Estimators for Discrete Estimation," *Biometrika*, 68, 301-309.

Table 1: Sample Statistics of Key Variables

	Public School Math Sample		Regression Sample	
	Mean	SD	Mean	SD
10th Grade Math Test Score	51.306	9.850	51.561	9.776
Assigned Daily Hours of Homework	0.643	0.392	0.640	0.385
Weekly Hours of Math Class	3.922	1.033	3.968	1.035
8th Grade Math Test Score	51.488	9.931	51.804	9.955
Mother's Education				
High School Dropout	0.133	0.340	0.133	0.340
High School	0.396	0.489	0.421	0.493
Junior College	0.136	0.343	0.135	0.342
College Less Than 4 Years	0.097	0.296	0.090	0.287
College Graduate	0.146	0.353	0.133	0.339
Master Degree	0.069	0.255	0.070	0.255
Ph.D., MD., etc	0.019	0.137	0.016	0.125
Family Size	4.606	1.400	4.564	1.337
Female	0.498	0.500	0.491	0.499
Race				
Others	0.042	0.202	0.032	0.178
Hispanic	0.085	0.280	0.074	0.262
Black	0.117	0.321	0.093	0.290
White	0.753	0.499	0.799	0.400
% of Teachers Holding a Graduate Degree	0.508	0.499	0.513	0.499
Teacher's Race				
Others	0.017	0.131	0.011	0.107
Hispanic	0.017	0.129	0.017	0.131
Black	0.050	0.218	0.040	0.196
White	0.914	0.279	0.930	0.254
Teacher's Evaluation of the Overall Class Achievement				
High Level	0.254	0.435	0.271	0.444
Average Level	0.410	0.491	0.408	0.491
Low Level	0.236	0.424	0.217	0.412
Widely Differing	0.099	0.299	0.101	0.302
Class Size	23.521	7.315	23.387	7.210
Number of Observations	6913		4036	

NOTES: Weighted summary statistics are reported. The variables are only a subset of those utilized in the analysis. The remainder are excluded in the interest of brevity. The full set of sample statistics are available upon request.

Table 2: Parametric Estimates of Homework on 10th Grade Math Test Scores

	Coefficient (Standard Error)							
Homework	4.195	3.565	0.971	1.039	1.174	1.164	0.528	1.220
	(0.576)	(0.496)	(0.200)	(0.228)	(0.230)	(0.230)	(0.214)	(0.421)
Homework Squared	…..	…..	…..	…..	…..	…..	…..	-0.193
								(0.099)
8th Grade Math Test Score	…..	…..	0.806	0.799	0.797	0.795	0.708	0.707
			(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.011)
Class Size	…..	…..	…..	…..	…..	0.044	0.025	0.025
						(0.015)	(0.014)	(0.014)
Weekly Hours of Math Class	…..	…..	…..	…..	…..	0.034	-0.003	-0.000
						(0.107)	(0.102)	(0.102)
R^2	0.029	0.238	0.769	0.815	0.817	0.817	0.834	0.834
Other Controls:								
Demographic and Family Characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	No	No	No	No	Yes	Yes	Yes	Yes
Class Characteristics	No	No	No	No	No	Yes	Yes	Yes
Teacher's Evaluation of the Overall Class Achievement	No	No	No	No	No	No	Yes	Yes

NOTE: White standard errors are reported in paranthesis. See text for definition of the variables.

Table 3: Nonparametric Estimates of Homework on 10th Grade Math Test Scores

	Coefficient (Standard Error)
Mean	0.566
	(0.354)
10%	-0.405
	(0.377)
20%	-0.080
	(0.507)
30%	0.131
	(0.326)
40%	0.299
	(0.338)
50%	0.486
	(0.408)
60%	0.679
	(0.329)
70%	0.893
	(0.380)
80%	1.175
	(0.668)
90%	1.655
	(0.597)
R^2	0.934

NOTES: Standard errors are obtained via bootstrapping. Estimates control for individual, family, teacher, classroom characteristcs and school fixed effects.

Table 4: Quartile Estimates for Selected Continuous Regressors

	Coefficient (Standard Error)			
	Mean	Q1	Q2	Q3
8th Grade Math Test Score	0.751	0.687	0.760	0.816
	(0.040)	(0.058)	(0.021)	(0.062)
Class Size	-0.002	-0.042	-0.008	0.040
	(0.041)	(0.033)	(0.024)	(0.038)
Weekly Hours of Math Class	-0.209	-0.500	-0.191	0.079
	(0.265)	(0.251)	(0.468)	(0.197)

NOTE: Standard errors are obtained via bootstrapping.

Table 5: Parametric/Nonparametric Estimates of Homework on 10th Grade Math Test Scores by Achievement Level

	Coefficient (Standard Error)			
	High Achievement	Average Achievement	Low Achievement	Widely Differing
Nonparametric Estimates				
Mean	0.770	0.213	0.789	0.903
	(0.360)	(0.881)	(0.431)	(0.792)
0.10	-0.085	-0.555	-0.177	-0.603
	(0.447)	(0.393)	(0.573)	(1.661)
0.20	0.161	-0.276	0.234	0.051
	(0.304)	(0.619)	(0.473)	(0.750)
0.30	0.297	-0.104	0.478	0.425
	(0.579)	(0.635)	(0.681)	(0.669)
0.40	0.458	0.022	0.649	0.660
	(0.743)	(0.558)	(0.487)	(1.348)
0.50	0.631	0.172	0.787	0.896
	(0.422)	(1.032)	(0.500)	(0.609)
0.60	0.858	0.300	0.914	1.129
	(0.423)	(0.458)	(0.513)	(0.998)
0.70	1.109	0.462	1.088	1.373
	(0.434)	(0.417)	(0.470)	(1.162)
0.80	1.373	0.683	1.289	1.751
	(0.546)	(0.534)	(0.437)	(0.893)
0.90	1.922	1.076	1.719	2.457
	(0.643)	(1.212)	(0.431)	(1.197)
Parametric Estimates				
Homework	2.025	0.353	2.272	1.253
	(0.559)	(0.433)	(1.339)	(1.665)
Homework Squared	…..	…..	-0.557	…..
			(0.272)	

NOTES: Standard errors are obtained via bootstrapping for the nonparametric estimates and White standard errors are reported for the parametric estimates. Estimates control for individual, family, teacher, classroom characteristics and school fixed effects.

Table 6: Quartile Estimates for Selected Continuous Regressors by Achievement Level

Coefficient
(Standard Error)

	Mean	Q1	Q2	Q3	Parametric
High Achievement					
8th Grade Math Test Score	0.663	0.609	0.663	0.712	0.635
	(0.032)	(0.024)	(0.042)	(0.030)	(0.021)
Class Size	-0.035	-0.057	-0.035	-0.011	-0.015
	(0.024)	(0.032)	(0.026)	(0.036)	(0.029)
Weekly Hours of Math Class	-0.051	-0.220	-0.038	0.126	-0.001
	(0.147)	(0.206)	(0.239)	(0.209)	(0.264)
Average Achievement					
8th Grade Math Test Score	0.782	0.738	0.779	0.822	0.715
	(0.027)	(0.023)	(0.049)	(0.027)	(0.019)
Class Size	-0.012	-0.038	-0.011	0.013	0.004
	(0.028)	(0.040)	(0.040)	(0.022)	(0.027)
Weekly Hours of Math Class	-0.512	-0.735	-0.492	-0.292	-0.245
	(0.248)	(0.264)	(0.213)	(0.190)	(0.183)
Low Achievement					
8th Grade Math Test Score	0.771	0.737	0.781	0.823	0.655
	(0.031)	(0.061)	(0.058)	(0.055)	(0.272)
Class Size	0.065	0.036	0.064	0.090	0.129
	(0.024)	(0.022)	(0.027)	(0.023)	(0.054)
Weekly Hours of Math Class	0.028	-0.172	0.017	0.243	0.349
	(0.184)	(0.215)	(0.253)	(0.216)	(0.311)
Widely Differing					
8th Grade Math Test Score	0.824	0.771	0.840	0.889	0.712
	(0.052)	(0.086)	(0.062)	(0.088)	(0.052)
Class Size	-0.009	-0.008	-0.004	0.056	0.028
	(0.054)	(0.102)	(0.058)	(0.062)	(0.073)
Weekly Hours of Math Class	0.038	-0.313	0.049	0.426	-0.209
	(0.328)	(0.248)	(0.325)	(0.352)	(0.697)

NOTE: Standard errors are obtained via bootstrapping for the nonparametric estimates and White standard errors are reported for the parametric estimates.

Table 7: Stochastic Dominance Tests of the Coefficient Distributions

	Equality of Distributions	First Order Stochastic Dominance
	p-values	p-values
High Achievement/Average Achievement	0.000	0.986
High Achievement/Widely Differing	0.000	0.000
Low Achievement/High Achievement	0.000	0.258
Low Achievement/Average Achievement	0.000	0.968
Low Achievement/Widely Differing	0.000	0.000
Widely Differing/Average Achievement	0.000	0.970

NOTES: Probability values are obtained via bootstrapping. The null hypothesis is rejected if the p-value is smaller than some significance level α ($0<\alpha<1/2$).

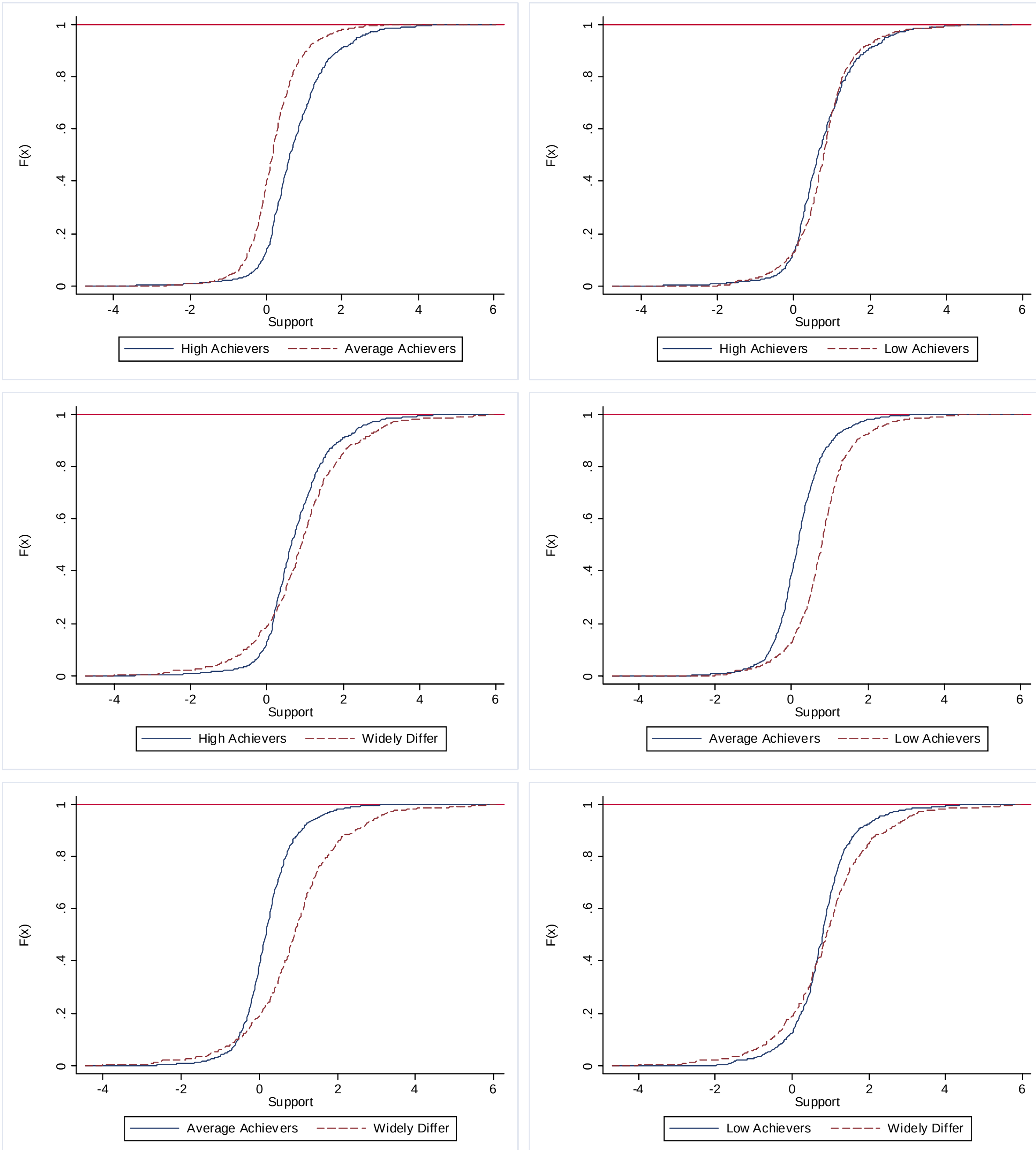

Figure 1: CDFs-Estimated Homework Coefficients by Achievement Level

Appendix A :

Table A1: Sample Statistics of Key Variables for Students with Negative Respsonses to Additional Homework

	Negative Response Sample	
	Mean	SD
10th Grade Math Test Score	51.829	9.018
Assigned Daily Hours of Homework	0.682	0.460
Weekly Hours of Math Class	3.875	1.054
8th Grade Math Test Score	51.447	9.010
Mother's Education		
High School Dropout	0.123	0.329
High School	0.494	0.500
Junior College	0.096	0.295
College Less Than 4 Years	0.073	0.260
College Graduate	0.104	0.306
Master Degree	0.085	0.279
Ph.D., MD., etc	0.022	0.148
Family Size	4.529	1.284
Female	0.576	0.494
Race		
Others	0.016	0.129
Hispanic	0.068	0.252
Black	0.119	0.324
White	0.795	0.403
% of Teachers Holding a Graduate Degree	0.576	0.494
Teacher's Race		
Others	0.007	0.087
Hispanic	0.023	0.151
Black	0.056	0.494
White	0.912	0.282
Teacher's Evaluation of the Overall Class Achievement		
High Level	0.161	0.367
Average Level	0.648	0.477
Low Level	0.121	0.327
Widely Differing	0.068	0.252
Class Size	24.005	6.562
Number of Observations	955	

NOTES: Weighted summary statistics are reported. The variables listed are only a subset of those utilized in the analysis. The remainder are excluded in the interest of brevity. The full set of sample statistics are available upon request.

Appendix B :

Table B1: Parametric/Nonparametric Mean Estimates of Homework on 10th Grade Math Scores By Gender/Race

	Coefficient (Standard Error)	
	Nonparametric	Parametric
Male	0.542	0.680
	(0.410)	(0.603)
Female	0.429	0.454
	(0.307)	(0.367)
Hispanic	0.832	0.624
	(0.768)	(0.476)
Black	1.178	0.523
	(1.443)	(0.601)
White	0.554	0.559
	(0.259)	(0.389)

NOTES: Standard errors are obtained via bootstrapping for the nonparametric estimates and White standard errors are reported for the nonparametric estimates. Estimates control for individual, family, classroom, teacher characteristics and school fixed effects.

Table B2 : Stochastic Dominance Tests of the Coefficient Distributions by Gender/Race

	Equality of Distributions	First Order Stochastic Dominance
	p-values	p-values
Male/Female	0.000	0.991
White/Hispanic	0.106	0.173
White/Black	0.081	0.128
Hispanic/Black	0.279	0.823

NOTES: Probability values are obtained via bootstrapping. The null hypothesis is rejected if the p-value is smaller than some significance level α ($0<\alpha<0.5$).

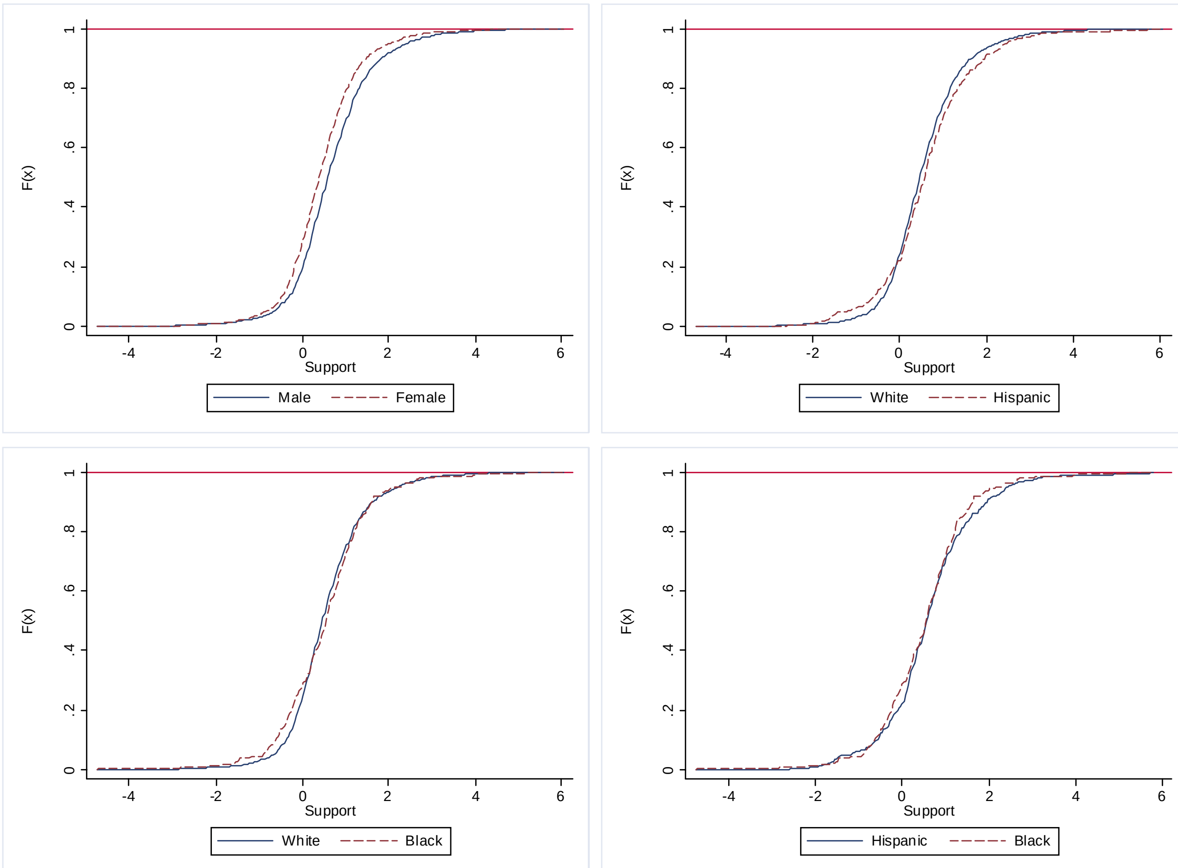

Figure B1: CDFs-Estimated Homework Coefficients by Gender/Race