ISERP
Working Paper 07-17

Teacher Effects on Academic and Social Outcomes in Elementary School

JENNIFER BOOHER-JENNINGS
Department of Sociology
Columbia University

THOMAS A. DIPRETE
Department of Sociology
Columbia University

Columbia University in the City of New York
December 2007

INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY
PIONEERING SOCIAL SCIENCE RESEARCH AND SHAPING PUBLIC POLICY

TEACHER EFFECTS ON ACADEMIC AND SOCIAL OUTCOMES IN ELEMENTARY SCHOOL*

Jennifer Booher-Jennings[1]
Department of Sociology
Columbia University

Thomas A. DiPrete[2]
Department of Sociology
Columbia University

ISERP Working Paper 07-17

December 2007

* This research was supported by grants from the Spencer Foundation and from the American Educational Research Association. The authors analyzed restricted data from the ECLS-K under terms of a license between DiPrete and the National Center for Education Statistics.
[1] Direct all correspondence to: Jennifer Booher-Jennings, Columbia University Department of Sociology, 1180 Amsterdam Avenue, 413 Fayerweather Hall, New York, NY, 10027. Email: jlj2102@columbia.edu
[2] Department of Sociology, 1180 Amsterdam Avenue, 401 Fayerweather Hall, Mail Code 2562, New York, NY 10027. Email: tad61@columbia.edu.

Teacher Effects on Academic and Social Outcomes in Elementary School

Abstract

Numerous studies conclude that teacher effects on academic achievement are substantial in size. Education is about more than academic achievement, and we know very little about teachers' effectiveness in promoting students' social development. Using data from the Early Childhood Longitudinal Study -- Kindergarten Cohort (ECLS- K), we estimate teacher effects on social as well as academic outcomes. We find that teacher effects on social development are sizeable, and are approximately twice as large as teacher effects on academic development. We further determine that teachers who produce better than average academic results are not the same teachers who produce better than average social results. However, we find that observable characteristics of teachers and the instructional approaches utilized in their classrooms are weak predictors of teacher effects. Finally, we show that the development of social skills has a positive effect on the growth of academic skills, and therefore teachers who are good at teaching social skills provide an additional indirect boost to academic skills in addition to their direct teaching of academic skills. We conclude that current policy debates over what it means to be a "highly qualified teacher" should also take social development into account.

After decades of searching for school effects, scholars have turned their attention to the classroom. Early studies of school effects (Coleman et al. 1966; Hauser, Sewell, and Alwin 1976; Jencks 1972) failed to separate schools, the organizations that conduct instruction, from schooling, the process through which instruction occurs (Bidwell and Kasarda 1980). Recent studies recognize that individual teachers direct and shape students' instructional experiences, and focus on the extent to which teachers differ in their ability to improve student achievement.

This burgeoning literature demonstrates that teacher effects on academic achievement are substantial in size (Clotfelter, Ladd, and Vigdor 2006; Murnane 1983; Rivkin, Hanushek, and Kain 2005; Rockoff 2004). However, this literature is limited in important respects. First, the intense focus on short-term academic outcomes leaves social development out of the picture. Social development is an important component of education both as an end in itself and as a probable determinant of long-term academic progress in school. Second, it is not known whether teachers can influence social development; if they can, it becomes of interest to understand whether the teachers who produce better than average academic results are the same teachers who produce better than average social results. Moreover, if social development is an important component of academic development, and if the production of social development requires different competencies than the production of math and reading development, then we miss an important social policy tool by leaving social development out of the study of teacher effects.

Using data from the first five waves of the Early Childhood Longitudinal Study--Kindergarten Cohort (ECLS-K), we build upon previous studies of teacher effects to

address both academic and social development. To motivate the importance of examining social development, we first determine the extent to which social development affects academic development. We then estimate the size of teacher effects on social development in order to understand how the size of these effects compares with teacher effects on academic development. We ask whether teacher competencies in fostering social development are tightly or loosely coupled with teacher competencies in fostering academic development. Finally, we assess the extent to which observable characteristics of teachers and instructional approaches account for differences in teacher effectiveness. In answering these questions, we inform current debates over what it means to be a "highly qualified teacher."

Literature Review

Compared to the effect sizes of other common measures of school quality, such as school resources, instructional interventions, and class size reductions, teacher effects are large (Odden, Borman, and Fermanich 2004). In their review of the literature, Nye, Konstantopoulos, and Hedges (2004) found that 7% to 21% of the variance in achievement gains results from differences in teacher effectiveness. In their own analysis of the Tennessee STAR experiment, they determined that moving a student from the 25th to the 75th percentile of teacher effectiveness would increase reading and math gains by .35 and .48 standard deviations, respectively. Rowan, Correnti, and Miller (2002) identified much larger effect sizes, ranging from .77 to .78 for reading gains, and .72 to .85 for math gains. Rivkin, Hanushek, and Kain (2005) found a one standard deviation increase in teacher effectiveness is associated with a lower-bound gain of .11 standard deviations for math achievement and .10 standard deviations for reading, while Rockoff

found an effect close to the lower bound estimate of Rivkin et al. The different size of estimated effects arises partly from differences in the grade under study and other data issues and partly from differences in the methodological strategy that is employed to address the problems of self-selection and sampling variability.

If teachers matter as much as these studies suggest, a critical question is to what extent a teacher's performance can be predicted by observable characteristics such as experience, education, certification, and test scores. Numerous studies conclude that experienced teachers are more effective in increasing student achievement (Clotfelter, Ladd, and Vigdor 2006; Greenwald, Hedges, and Laine 1996; Murnane 1983; Rivkin, Hanushek, and Kain 2005; Rockoff 2004), with effect sizes ranging from .04 to .13. In a particularly comprehensive treatment of teacher experience, Clotfelter, Ladd, and Vigdor (2006) found that having a highly experienced teacher in the fifth grade– that is, a teacher with more than 27 as compared to zero years of experience – is associated with an increase of .13 standard deviations for math and .095 standard deviations for reading, with half the gain occurring in the first two years of teaching. Other studies find that measures of teachers' ability, as captured by standardized tests or licensure scores are positively associated with student achievement (Ferguson 1991, 1998; Ferguson and Ladd 1996; Clotfelter, Ladd, and Vigdor 2006; Rowan, Chiang, and Miller 1997). Most of the variation in teacher quality, however, is not captured by observed teacher characteristics.

Taken together, the existing studies have greatly improved our knowledge of teachers' effects on student academic achievement. Education is about more than academic achievement, however, and we know very little about schools' or teachers'

effectiveness in achieving other educational goals. In particular, little is known about the determinants of social development, including a positive orientation to learning, the ability to interact in a positive way with teachers and other students, or the ability to observe school rules and avoid disruptive behaviors such as fighting with other students. It is possible that teachers that are effective in promoting academic growth also promote students' social growth. On the other hand, these teacher qualities may be largely independent of each other, whether because they call on different abilities, or because they have not been emphasized to the same extent in teacher education. It may even be the case that instruction in academic and social skills may compete with each other, with the consequence that specific teachers excel in either one area or the other.

We have identified only two studies (Alexander, Entwisle, and Thompson 1987; Downey and Pribesh 2004) that address the relationship between teachers' attributes and students' social outcomes. Neither of these studies, however, specifically estimates teacher effects on social outcomes. Rather, they both address how student-teacher status differentials (measured in terms of class or race) result in low status students' receipt of poorer evaluations. Downey and Pribesh (2004) used nationally representative samples of kindergartners (the Early Childhood Longitudinal Study – Kindergarten Cohort) and adolescents (the National Educational Longitudinal Study) to examine the effects of student-teacher racial matching on teachers' evaluations of students' externalizing problem behaviors and approaches to learning. They found that black students receive poorer behavioral ratings when they are matched to white teachers, with effect sizes of .05 to .1 standard deviations. Because Downey and Pribesh were interested in perceptions, they used cross-sectional teacher evaluations as their dependent variable. A

study of teacher effects on social outcomes, as opposed to teacher effects on perceptions, would require a control for students' initial position and better attention to the accurate measurement of student social development.

Alexander, Entwisle, and Thompson (1987) examined the effects of teacher-student social background matching in the first grade on teachers' evaluations of students' maturity. The authors found that students' race strongly conditioned the evaluations of teachers from high status backgrounds, but had no effect on the evaluations of low SES teachers. Like Downey and Pribesh, Alexander et al. do not address change in teachers' ratings of students over time.

In sum, the current literature leaves unaddressed the questions of social development's effects on academic development and the impact of teacher quality on social development. Our paper uses data from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K) to fill this gap in several respects. First, we identify a set of dimensions of social development and establish their impact on later academic achievement. Second, we estimate the impact of kindergarten teacher quality on social development and compare these effects with published estimates of the size of teacher effects on mathematics and reading achievement. In constructing these estimates, we use a variety of strategies (including the use of social development ratings by parents) that address the potential bias in these estimates that stems from the fact that social development is measured by teacher ratings. Third, we estimate the correlation between teacher quality in social development and teacher quality in academic achievement in order to determine whether these teaching skills are tightly coupled in the current population of teachers. Fourth, we use growth curve models to estimate the impact of

social development on subsequent academic development. Finally, we combine our estimates of teacher effects with our estimates from the growth curve models to estimate the indirect effects of teachers on academic achievement that operate through their impact on social development.

We focus our attention on teachers in early education because of our theoretical expectations that teacher effects on social development are likely to be larger for younger children than for older children. This expectation stems from the broader literature on social development, which finds that social behaviors are most plastic in early childhood (Campbell et al. 2002, Hawkins et al. 2001, 2005; MacDonald 1985; Nelson 1999; Stiles 2000; Yoshikawa 1995). Little is known about social development in the education process, however, and it is possible (and indeed, we hope) that teachers can shape a student's social behaviors in a positive way throughout elementary school and into high school. We therefore see our paper as a starting point for a broader effort that focuses on multiple points in the educational process.

Data and Methods

The ECLS-K is a study of a nationally representative sample of 21,260 kindergarteners who attended kindergarten in the 1998-1999 school year, and who have now been followed through fifth grade.[1] These data provide parent reports on the socioeconomic and demographic characteristics of the children, teacher and parent reports of their social development, cognitive assessments, and measures of teacher and school characteristics. The ECLS-K was designed as a multilevel study that collected data on multiple kindergarten children for the same teacher, often for multiple classrooms

in the same school.[2] This multilevel character allows us to estimate the effects of teacher quality on student development.

The estimation of teacher effects is complicated by the problem of non-random selection, and the strategies used in contemporary research differ in part because of the strengths and limitations of the alternative data sets. In order to evaluate the strategy allowed by ECLS-K, we need to place it in the context of recent methodological strategies employed by scholars who have estimated teacher effects on academic outcomes. Rivkin, Hanushek and Kain (2005) had data on test scores across multiple grades for three cohorts of students in Texas, but lacked information on the specific identity of the teachers. By computing changes across grade for specific cohorts, they were able to eliminate the unmeasured fixed effects of students and families. They computed the difference in these differences across the same grades for different cohorts and attributed the difference to the change in the mix of teachers over time. Using information on teacher turnover along with a set of strong assumptions (that teacher exit is exogenous, that a teacher is equally effective across cohorts, and that there is no measurement error in the cohort data), they were able to estimate a lower bound on the teacher effect.

Clotfelter, Ladd, and Vigdor (2006) used administrative data for all North Carolina elementary students. They examined whether elementary students in the same school but in different classrooms were statistically distinguishable across six criteria and grouped schools in to the 45% where the students were not distinguishable and the 55% where they were distinguishable on at least one criterion. For both groups of schools they estimated fixed effects models for teachers both including and omitting lagged test

scores. Under the theory of random assignment, the estimated teacher effects should not vary across the two specifications or between the two sub-samples of schools, and they found this to be true when they included fixed effects for schools and an extensive set of student controls in the model.

Rockoff (2004) estimated teacher effects on academic outcomes using data for two New Jersey school districts that linked teachers with students who were followed for up to twelve years. The ability to observe the same teachers across multiple cohorts allowed Rockoff to estimate multiple teacher effects for the same teacher and thereby separate the "permanent" teacher effect from transitory effects that were due in part to sampling variability on student outcomes within classrooms. He found that the variation in "permanent" effects, while substantively important, was only about half of the variation estimated for any given year.

The ECLS-K data have the advantage over these other sources of providing detailed measures of social development and provide data on schools that have more than one kindergarten teacher and therefore allow the estimation of models that control for the nonrandom selection of students into schools. Like the North Carolina data the ECLS-K data contain detailed student controls and therefore – at least supported by the results of Clotfelter et al (2006) – we can adequately control for non-random assignment to students to classrooms within schools. The ECLS-K data have the disadvantage of being able to estimate teacher effects for only one cohort, which – based on the Rockoff results – would lead to an overestimate of teacher effects. We address this issue by employing a set of alternative estimation strategies that include conservative estimates of the estimates of teacher effects on growth in social skills between the start and end of kindergarten.

In order to separate school and teacher effects, we restricted our sample to first-time kindergarteners attending schools with two or more sampled kindergarten teachers. Furthermore, to accurately estimate teacher effects, we further restricted our sample to include only students in classrooms with three or more sampled students. Of the originally sampled 21,409 kindergarten students, we excluded 10,640 observations because they lacked measures of academic or social skills in kindergarten or first grade, 420 observations because these students were not enrolled in kindergarten for the first time, 660 observations because of missing covariates, 1,356 observations because students were in classes with fewer than three students[3], and 2,720 students in schools with only one sampled teacher. In our analyses of kindergarten and first grade data, our final sample included 5,613 children taught by 1,093 teachers in 439 schools. In our analyses of growth from kindergarten through third grade, our sample included 4,792 children for reading and 4,814 children for math.

Measures of Academic and Social Skills: Dependent Variables

Our analyses make use of students' test scores in reading and math at the beginning and end of kindergarten, at the end of first grade, and at the end of third grade. The ECLS tests were designed to reduce ceiling and floor effects. To this end, students were first administered a routing test which determined the level of difficulty of their subsequent test. The ECLS then employed item response theory to place students on a common 64 point scale for mathematics and 92 point scale for reading. To ease interpretation, we converted these scores to percentile units.[4]

Teachers were asked to rate student social development at the beginning and end of kindergarten, the end of first grade, and the end of third grade. In order to identify the

major dimensions underlying the five social scales that are available in ECLS-K, we conducted an exploratory factor analysis.[5] Three scales – the Approaches to Learning scale, the Self-Control scale, and the Interpersonal skills scale – loaded primarily on one factor. The Approaches to Learning Scale rates the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. The Self-Control Scale indicates the child's ability to control behavior by respecting the property rights of others, controlling temper, accepting peer ideas for group activities, and responding appropriately to pressure from peers. The Interpersonal Skills scale rates the child's skill in forming and maintaining friendships, getting along with people who are different, comforting or helping other children, expressing feelings, ideas and opinions in positive ways, and showing sensitivity to the feelings of others (NCES 1999). The loadings for this factor analysis are displayed in Table A1. Hereafter, we refer to this dimension of social skills as learning-related/interpersonal skills.

Because the remaining two social scales formed separate dimensions in the factor analysis, we analyzed each of them separately. The Externalizing Problem Behaviors scale includes acting out behaviors such as the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities. The Internalizing Problem Behaviors Scale rates the student on the apparent presence of anxiety, low self-esteem, loneliness, and sadness. We reverse coded these two scales to remain parallel with our first measure; that is, moving up the scale implies positive social development. Receiving a higher rating on these indicators means that a student exhibited *fewer* externalizing or internalizing problem behaviors.

Independent Variables

Because students are not randomly assigned to teachers, the issue of selection bias must be addressed in any study of teacher effects. We control for variables that have been associated with students' academic and social development in previous research. These variables include race, gender, socioeconomic status, family structure, the presence of a biological mother, whether the student is an only child, home language, disability status, the student's age, AFDC receipt, and whether the student attends a full-day kindergarten. Descriptive statistics for these variables can be found in Table 1.

In the final section of this paper, we examine the extent to which instructional characteristics are associated with our estimates of teacher effects. Following Milesi and Gamoran (2006), we constructed four instructional scales, which capture the range of curricular approaches used to teach reading and math in elementary classrooms. To construct these scales, we summed multiple items. The first scale, the "whole language" scale, captures the frequency with which students write words with invented spellings, write stories/reports, write in a journal, and choose books the read. The second scale, the "phonics" scale, includes the frequency with which students work on letter names, practice writing the alphabet, work on phonics, and work on workbooks and worksheets. The third scale, the "teaching for understanding of math" scale, includes the frequency with which students work with counting manipulatives, solve math problems in small groups or with a partner, and work on problems that reflect real-life situations. The final scale, the "math drill" scale, captures the frequency with which students do math worksheets, use math textbooks, and do math on the chalkboard. Descriptive statistics for these scales can also be found in Table 1.

Analytic Strategy

Our study includes four components. We first estimate academic growth from kindergarten through third grade in order to establish the impact of social development on academic development. We then estimate teacher effects on social skills and compare them with estimated teacher effects on academic outcomes using alternative methods in order to establish the robustness of our results. Third, we determine whether the teachers who are good at academic outcomes are the same teachers as those who are good at social outcomes. Finally, we decompose teacher effects on academic outcomes in later elementary grades in order to determine the indirect importance of being a good *social skills* teacher on subsequent *academic* development.

We begin with a multilevel model of academic growth where academic scores (level 1) are nested within students (level 2). This model takes the form:

$$\begin{aligned} y_{it} = {} & \beta_1 + \beta_2 \text{AGE}_{it} + \beta_2 \text{AGE}_{it}^2 + \beta_3 \text{SCORE}_{i,t-1} + \beta_4 \text{RACE}_i \\ & + \beta_5 \text{SINGLE PARENT FAMILY}_i + \beta_6 \text{SES}_i + \beta_7 \text{FEMALE}_i \\ & + \beta_8 \text{SOCIAL SKILLS}_{i,t-1} + \zeta_i^{(1)} + \zeta_i^{(2)} \text{ AGE}_{ij} + \varepsilon_{ij} \end{aligned} \qquad (1)$$

where y_{it} is the score of student i at time t, $\zeta_i^{(1)}$ is a random intercept that varies across students, and $\zeta_j^{(2)}$ is a random slope that allows students to vary in their rate of growth. Because we include a measure of students' academic position (SCORE) at time i-1, we estimate growth between the beginning and end of kindergarten, the end of kindergarten and the end of first grade, and the end of first grade and the end of third grade. RACE is a vector of dummy variables, where the reference category is white. Because approximately half of our sample changes schools between kindergarten and third grade, we do not include a school-level random effect. We first estimate three separate models for both math and reading scores, where social skills represent respectively learning-

related/interpersonal skills, externalizing problem behaviors, and internalizing problem behaviors. While this represents a simplification of reality (students simultaneously possess social skills across all three domains), it clarifies the size of the effect of social skills on growth in academic skills. To test for non-linear effects of social skills on academic growth, we then re-estimate this growth model, and include dummy variables representing the quartile of the social skills distribution in which the student is situated.

In the second part of our study, we estimate teacher effects on academic and social outcomes. The measurement of teacher effects on social development is complicated by two main issues. The first has to do with dimensional structure. Social development in ECLS-K can be conceptualized along more dimensions than academic development, and these dimensions are less-well specified. The second issue concerns measurement bias. While a standardized testing instrument evaluated all students in reading and math, teachers rate their own students' social skills, and this fact makes it difficult to distinguish differences in kindergarten teachers that are due to objective differences in social development of students from differences in how kindergarten teachers rate their students. We address this problem by constructing measures of social development that do not depend on the ratings of the kindergarten teacher, and compare these with measures based on the kindergarten teacher ratings.

Specifically, we adopt three different methods of measuring teacher effects on social and academic development. Method 1 is our baseline in which we do not take into account the impact of possible ratings bias in our estimates of teacher effects on social development. Our dependent variables for method 1 are the kindergarten teacher's ratings of the students' social skills and math and reading test scores at the end of

kindergarten. We control in these models for teachers' social rating at the beginning of kindergarten, and students' test scores at the beginning of kindergarten. In Method 2, we use academic and social measures at the end of first grade in order to address the endogeneity that arises from using kindergarten teachers' own ratings to assess their effectiveness in promoting growth in social development, but we continue to control for kindergarten teachers' ratings of students' social skills at the beginning of kindergarten and students' test scores at the beginning of kindergarten. In Method 3, we drop the use of kindergarten teacher ratings at both the origin and the destination time point. As in method 2, our dependent variables are first grade teachers' social ratings and test scores at the end of first grade. However, we create a new baseline measure of social development. Specifically, we regress the kindergarten teacher's social ratings on a series of predictor variables from the parent survey, and we include the predicted value of the kindergarten teachers' rating from this regression as our measure of social skills at the origin time (see Table A2). We use each of these three methods in the analytic approaches discussed in the balance of the methods' section.

To obtain estimates of teacher effects, we separately estimate random effects and fixed effects models. The random effects model is a three-level hierarchical linear models, where students (level 1) are nested within teachers (level 2), who are nested within schools (level 3), i.e.,:

$$y_{ijkt} = \boldsymbol{\beta}'\mathbf{X}_{it} + \zeta_{jk}^{(2)} + \zeta_{k}^{(3)} + \varepsilon_{ijkt} \qquad (2)$$

where y_{ijkt} is a measure of a student's achievement at times t, i is the child in the classroom of teacher j in school k, $\mathbf{X_{it}}$ are characteristics of the child and the child's family including the score or rating at *time 1*, race, gender, socioeconomic status, family structure, the child's age, the presence of the biological mother, whether a language

besides English is spoken at home, student disability, AFDC receipt, full-day kindergarten. In the models where scores or ratings at the end of first grade are the dependent variables, we also include student retention in kindergarten and whether the student has the same teacher again. β are the fixed parameters. The random intercept $\zeta_{jk}^{(2)}$ varies across teachers and therefore schools, while the random intercept $\zeta_{k}^{(3)}$ varies across schools. Using these results we calculate the intraclass correlation for the teacher and the school and compare the results for social and academic outcomes using the formulas:

$$\rho(teacher\ within\ schools) = corr(y_{ijk}, y_{i'jk} \mid j) = \frac{\psi^{(2)}}{\psi^{(2)} + \theta}$$
$$\rho(school) = corr(y_{ijk}, y_{i'j''k} \mid school\ characteristics) = \frac{\psi^{(3)}}{\psi^{(2)} + \psi^{(3)} + \theta} \quad (3)$$

In order to explore the factors that covary with teaching effectiveness, we compute empirical Bayes estimates of teacher effectiveness. We use these estimates to assess whether good teaching is a general skill that implies positive outcomes across the range of student achievement dimensions, or whether it involves specialized and at least to some extent independent competencies. The empirical Bayes estimate of teacher effectiveness is the mean of the posterior distribution of $\zeta^{(2)}$, and the variance of the prediction errors of $\zeta^{(2)}$ depends on the number of students observed per teacher, such that:

$$Var\left(\hat{\zeta}_{jk}^{(2)EB} - \zeta_{jk}^{(2)}\right) = \frac{\hat{\theta}/n_{jk}}{\hat{\psi}^{(2)} + \hat{\theta}/n_{jk}}\hat{\psi}^{(2)} \quad (4)$$

where n_{jk} is the number of student respondents in the classroom of teacher j in school k, $\hat{\psi}^{(2)}$ is the estimated variance of the random teacher effects, and $\hat{\theta}$ ís the estimated error

variance. At the individual-level, of course, these estimates of quality – being based upon outcomes involving as few as three students – are not measured with high reliability. Our interest instead is to use these estimates to estimate the spread of competence across the distribution of estimated teacher effects. We also use these estimates to study the association between teacher effectiveness and the four instructional scales described above.

The second method is a modified version of the fixed-effects estimator proposed by Nye, Konstantopoulos, and Hedges 2004 in their meta-analysis of teacher effects studies. This approach compares the R^2 when prior achievement, demographic variables, and school effects are controlled with the R^2 when teacher as well as school dummy variables are included. Nye et al. argued that ΔR ($\sqrt{(R^2_2 - R^2_1}$)), can then be "loosely" interpreted as a standardized regression coefficient. Nye et al.'s method probably overstates the size of teacher effects because it does not account for the multiple degrees of freedom that are used up when the teacher dummy variables are added to the equation. We modified their proposed method by using adjusted R^2 in the computation.

Results

We begin by examining the distributions of ECLS-K's measures of academic and social skills. These distributions are displayed in Figure 1. Panels A and B demonstrate that the ECLS-K reading and math assessments did reduce ceiling effects; the right skew in these distributions shows that these assessments distinguished various degrees of high achievement. In contrast, the left skew in the distributions shown in Panels C, D, and E suggests that the three measures of social skills do a better job of distinguishing poor behavior than gradations of good behavior.

Does social development affect academic development?

While social development is important as an end in itself, we are also interested in the extent to which social development affects academic development. In order to estimate the cross-over effects between social and intellectual development, we estimate growth-curve models of academic development in which the rate of academic growth is specified to depend upon social development. Estimates of these effects in combination with estimates of the impact of teachers on social development allow a determination of the "total" effect of a teacher as the combined "direct" effect on academic growth plus the "indirect" effect on academic growth via the teacher's impact on social development. This conceptual model is illustrated in Figure 2 (while we posit the existence of the dashed arrow in Figure 2, we do not estimate this model in the current paper).

Table 2a displays the results of these models. For each of our three measures, social skills have a positive effect on math and reading growth. The effect sizes of social skills on reading and math growth are almost identical. Learning-related/interpersonal social development has a larger effect on academic skills than does the component of social development that reduces externalizing or internalizing problem behaviors. For reading, an increase of 10 percentile points in learning/interpersonal social development increases reading scores by .62 percentile points. An increase of ten percentile points in the externalizing behaviors scale increases reading scores by .39 percentile points, while the same increase in internalizing behaviors increases reading scores by .23 percentile points.[6]

In Table 2b, we estimate the same model, but instead divide each of the three social measures into quartiles, where the fourth quartile, for students with the highest level of social development, is the reference category. The results in Model 1 suggest

that learning/interpersonal skills has non-linear effects on academic development. Students gain only a small benefit (.733 percentile points for math and 1.191 percentile points for reading) from being in the fourth rather than the third quartile. However, the penalty for second quartile relative to third quartile location is a larger 1.93 points for math and 1.946 points for reading. Being in the first quartile relative to the third produces an additional 2.035 point penalty for math, and a 1.056 point penalty for reading.

In Model 2, we add dummy variables for quartile 1 of externalizing and internalizing problem behaviors. We find that being in the first quartile of externalizing and internalizing problem behaviors has no statistically significant effect on reading or math growth. In Model 3, we add a dummy variable coded as 1 if a student is in quartile one for all three measures of social skills. The coefficient on this interaction does not reach statistical significance.

Tables 2a and 2b show that social development has a direct effect on academic growth. Because learning-related/interpersonal skills have stronger effects on academic development than the remaining two measures of social skills, the balance of this paper focuses on estimating teachers' impacts on these skills.

How large are teacher effects on academic and social development?

We next examine in Table 3a how much of the variation in social and academic outcomes lies between schools and between kindergarten teachers within schools. Using equation (2), we obtained intraclass correlations (ICC) for an unconditional model, which contained no measured covariates, and then with a model that included control variables

for the classroom students. We report the coefficient estimates for this model in appendix table A3. Here we focus on the resulting intraclass correlations.

Beginning with the unconditional models for Method 1 (which use kindergarten teacher ratings as the outcome measure for social development), we find that .219 and .187 of the student variance is between schools for reading and math outcomes, respectively, a fact that we attribute largely to nonrandom selection. The proportion of social outcome variance that lies between schools is smaller, ranging from .037 to .042. We believe that the difference in the school variance for social and academic outcomes suggests that teachers within schools rate students on social outcomes in comparison with their school peers rather than with the broader population of students across the country.

The between-teacher variance for reading and math (.051 and .027) in the unconditional models is much smaller for academic skills than for social skills, where the teacher variance ranged from .125 to .268. When we control for socioeconomic, demographic, and prior performance covariates to address the non-random assignment of students to schools and classrooms, we find that the between-teacher variance for academic outcomes remains smaller than the between-teacher variance for social outcomes. By comparing the between-teacher variance on social effects across the three measurement methods, it is clear that the kindergarten teachers' self-ratings of kindergarten student social skills is –as expected—upwardly bias our estimates of teacher effects on social skills. Method 3 does not utilize kindergarten teacher ratings for either the starting or the ending measurement of social development, and – as we expected – the ICC that is between teachers but within schools is dramatically smaller than for methods

2 or 1. Importantly, the method 3 results demonstrate that teacher effects on social outcomes are at least as large as teacher effects on academic outcomes.

These differences can similarly be observed if we contrast teacher effects at the 25th and 75th points in the teacher effects distribution (see Table 4). Using Method 1, which is the most accurate for estimating kindergarten teacher effects on *academic* achievement, moving a student from the 25th percentile to the 75th percentile of the teacher effects distribution would increase achievement by .17 standard deviations for math and .27 standard deviations for reading. The size of these effects is smaller than those established by Nye et al (2004), who found that moving a student from the 25th to the 75th percentile of the distribution would increase math test scores by .48 standard deviations and reading test scores by .35 standard deviations. Older studies, such as Armour (1976) and Hanushek (1992), find effects similar to those in Nye et al. In the Armour study, which included primarily African-American and Latino students in Los Angeles, a 25-75th percentile shift in teacher effectiveness yields a gain of .35-.50 standard deviations for reading; in Hanushek's study, this shift produces a gain of .43 standard deviations for reading. We expected that kindergarten teachers would have smaller effects on academic achievement than other elementary grade teachers, and suggest that the effect size differences described above are in part explained by the grade under study.

To compare academic and social teaching effects, we focus on the method 3 results in Table 4. The results in Table 4 support the results in table 3 in demonstrating that social teaching effects are larger than academic teaching effects for kindergarten teachers. Moving a student from a below-average to an above-average kindergarten

teacher would increase student achievement by approximately 2.099 percentile points in math and 3.636 percentile points in reading between the start of kindergarten and the end of first grade, and 5.283 percentile points in learning-related/interpersonal social skills.

Rockoff showed that the estimation of teacher effects based on single-classroom data overestimates the teacher effect variance by about 100%. Table 4 shows that the use of methods 2/3, which contrast test scores at the end of first grade with scores at the beginning of kindergarten, understates the estimated teacher effects obtained from method 1 for math and reading by about 50%. Thus, methods 2/3 appear to give an estimate of teacher effects on academic growth from the beginning to the end of kindergarten that would approximate what we might have obtained if we could estimate teacher effects across the same teacher with multiple cohorts of students. This fact suggests that the mean method 3 estimated effects for social skill teaching might also therefore be a reasonable approximation to the mean estimated teacher effects from the beginning to the end of kindergarten that would be obtained with uncontaminated measures for the same teacher across multiple cohorts.

One possible reason why the estimated teacher effects on social development are larger than the estimated teacher effects on academic development is that social skills may not be measured as reliably as academic skills. Table A4 in Appendix A shows that the correlations over time in teacher ratings of social skills are somewhat lower than are correlations of academic test scores. In Appendix B, we discuss our estimates of the impact of possible differences in reliability of these measurements on estimated teacher effects. These results show that the differences in reliability magnify the estimated difference in teacher effects between social and academic skills. However, the impact of

reliability differences is less than our estimated difference in teacher effects from equation (1), which supports the conclusion that kindergarten teachers differ more in their ability to affect social development than they do in their ability to affect growth in math and reading scores.

Our estimates above were based on random effects models. As a robustness check, we then used the method of Nye et al. (2004) to estimate the distribution of fixed effects for teacher academic and social skill teaching abilities. These results are presented in Table 3b. Column 1 of this table shows the R^2 for a model with individual covariates and school fixed effects. Column 2 adds teacher fixed effects. Column 3 shows Nye et al.'s estimate of the effect of a standard deviation increase in teacher quality on student outcomes. For a benchmark, note that Nye et al. obtained a value of .32 for the Hanushek (1992) study of reading change in second through sixth grades, while we obtain a value of .224 for reading change from the beginning to the end of kindergarten with the ECLS-K data. This may indicate that kindergarten teachers have less of an impact on improvement in reading than do elementary school teachers, which is consistent with the fact that reading is generally taught more intensively in the elementary school grades than in kindergarten. As noted earlier, the Nye et al. method inflates estimates by failing to account for the change in R^2 due to the degrees of freedom used up in the addition of teacher dummy variables to the model. This is corrected in column 6. We focus attention on method 3, which corrects for the rating bias on social skills. The .171 result for social development teaching skills is slightly larger than Rockoff and Rivkin et al.'s estimates of the impact of teachers on math and reading in

higher grades, and – consistent with our random effects results – is larger than the estimates we obtain for the impact of kindergarten teachers on academic skills.

How tightly coupled are teacher competencies?

To determine how tightly coupled teacher competencies are, we produced empirical Bayes estimates of teacher effectiveness separately for reading, math, and social skills and then examined their correlations, which are reported in Table 5. These correlations range from .338 to .364. Using Method 1, we find a weak positive math-social correlation of .020, and a reading-social correlation of .081. Using Method 2, we find a weak negative math-social correlation of .131 and a reading-social correlation of .122. These correlations are almost certainly downwardly biased because they are based on relatively few students per teacher. Nonetheless, these results suggest that math and reading competencies are more weakly correlated with social competencies than they are with each other.

Do observable characteristics of teachers or instructional approaches predict teacher effects?

A key question for policymakers is the extent to which observable characteristics such as experience, education, and certification predict teacher effects. We therefore re-estimated equation (1) while including observed teacher characteristics in order to establish the extent to which differences in these characteristics can account for our estimated differences in teacher quality. These characteristics include dummy variables for teacher age (where the reference category is less than 35 years old), teacher experience (where the reference category is teachers with more than five years experience), a dummy variable coded as 1 if the teacher holds a Masters degree, a dummy variable coded as 1 if the teacher holds the highest certification available, and

dummy variables for teacher race and teacher*student race interactions. Table 6 shows that only receipt of the highest certification and having one year of experience exhibit statistically significant relationships with student social development outcomes. Our findings thus mirror previous research, which also finds that observable characteristics of teachers are weak predictors of student outcomes.

Measures of education and experience are plausibly exogenous determinants of teacher quality. In contrast, differences in instructional style may be as much a consequence as a cause of differences in teacher quality. This problem notwithstanding, it is of interest to establish whether instructional styles can statistically account for the observed variation in the quality of teaching social skills. We therefore, estimated a series of OLS regressions where the dependent variables were the empirical Bayes estimates of academic and social teacher effects that control for all teacher characteristics listed above, and the independent variables were scales of the frequency with which teachers used four instructional approaches: whole language and phonics in reading, and teaching for understanding and drill for math. Our results (see Table 7) show that instructional approaches do not predict teachers' effectiveness in promoting social development. While teacher effects on social development are substantial, they are not predicted either by observable attributes of teacher or by the instructional approaches they use in their classrooms.

What is the total impact of social development on academic development?

The first section of this paper estimated the effect of social development on academic growth, while the second section of this paper estimated the effect of teachers on social development. If teachers can affect social development, which itself has longer-

run implications for academic development, it therefore follows that the overall impact of teachers on academic development has two components, a direct effect on academic achievement, and an indirect effect which operates through social development (see Figure 1). To get a rough estimate of the size of the indirect effect of teacher quality on academic growth which runs through a teacher's impact on social development, we used the following procedure. We start with a version of equation (1) in which we estimate the effect of social development at the end of kindergarten on academic achievement in third grade, controlling for academic achievement at the start of kindergarten. We then multiplied these estimates by a change in a social development percentile score that we estimate would be obtained by having a teacher in the 75th percentile as opposed to the 25th percentile. The result is an estimate of the indirect impact of having a "good" as opposed to a "bad" teacher of social skills on math and reading achievement gains by third grade. If we further assume that first and second grade teachers also have a likely impact on social development, we can ask the question of how big an impact on reading or math achievement would one gain by having two good teachers of social skills in comparison to two bad teachers. If we use our most conservative method -- i.e., method 3 -- the gain in math percentile points that could be expected by moving a student from the 25th to the 75th percentile of the teacher distribution for social skills teaching is .241 percentile points for math, and .302 percentile points for reading. While hardly large, this effect actually exceeds the estimated direct effect of kindergarten teacher academic skills on third grade academic outcomes. We conclude that the indirect effects on math and reading gain that come from the teaching of social skills are potentially important both from a substantive and from a policy-making perspective.

Discussion

The central contribution of this study is that teachers vary widely in their ability to promote social development. Teacher effects on social development in kindergarten are substantial in size, and are at least twice as large as teacher effects on academic development. In our most conservative estimates, having a kindergarten teacher at the 75th percentile of the teacher effects distribution as opposed to one at the 25th percentile increases learning-interpersonal social development by 5.283 percentile points, which was larger than our estimates of teacher effects on academic skills.

In addition to establishing the effects of teachers on social development, this article provides strong evidence that students' social development influences their academic development. We determined that these effects are non-linear, such that moving a student beyond the third quartile of social development provides only minimal academic benefits. We propose two mechanisms to account for this finding. First, students may need a minimum level of social development before they can take full advantage of the academic environment. Beyond this threshold, it appears that further social development leads to relatively minor increases in academic development. Second, as suggested by Alexander, Entwisle, and Thompson (1987), students with higher levels of social development may benefit from closer relationships with their teachers, which may in turn spur academic development. Future research should attempt to test for the existence of these mechanisms. The fact that social development likely influences later academic development exposes a new pathway by which teacher skills can affect student outcomes.

Finally, we found that the teachers who are good at promoting social development may not be the same teachers that are good at promoting academic development. In

future research, sociologists of education should problematize the dominant unidimensional notion of teacher quality, which assumes that a "good" teacher is effective across all educational domains. That we find that most teachers essentially specialize in academic or social development suggests that the question we should be asking is "good at what?" Though public education has many goals, both research and policy have focused narrowly on measuring and promoting academic outcomes. Because social development provides a pathway to academic development and is an important end of education in itself, it needs to be integrated into research and policy agendas.

We must also ask whether the effects of social development on academic development are larger for some groups than others. It is possible that these effects are heterogeneous across the categories of gender, race, and class. Gaps in social development may thus contribute to gender, race, and class gaps in academic achievement, and understanding this relationship is an important area for future research.

Our study has two limitations that should be addressed in future research. The first limitation derives from the measurement of social effects in ECLS. Ideally, a study of social development would utilize the same raters to measure students' social skills. In the absence of such a measure, we estimated teacher effects on social development using both "biased" (Method 1 and 2) and "unbiased" (Method 3) measures. Though we focused our discussion on the estimates from Method 3 and erred on the side of underestimating teacher effects on social development, the true effect probably lies between the estimates generated from Method 1 and Method 3. We hope that future work can specify the magnitude of these effects more precisely. Nonetheless, we believe our results represent an important contribution to understanding the multiple ways that

teachers affect their students. A second limitation concerns the lack of sufficient data on multiple students per classroom in grades beyond kindergarten. Because we restrict our study to early elementary education, we do not address how teacher effects on social development change as students move through upper grades. Psychological studies of social development have found that social behaviors are more difficult to change as children age (Campbell et al. 2002, Hawkins et al. 2001, 2005; MacDonald 1985; Nelson 1999; Stiles 2000; Yoshikawa 1995). These studies suggest that teacher effects on social development may be attenuated in upper grades, but research is needed to establish the extent of attenuation, and specifically the possibility of specific interventions to support the social development of at-risk adolescents.

We conclude that current policy debates over what it means to be a "highly qualified teacher" should also take social development into account. The federal No Child Left Behind Act requires states to measure the extent to which schools employ highly qualified teachers and to implement strategies to guarantee that all students have access to such teachers. The law assumes that “highly qualified” teachers are equally effective across all educational domains. This study shows that teachers effective in promoting academic development are often less effective in promoting social development. In the future, these policy tradeoffs should be acknowledged.

Notes

[1] Because the fifth-grade data have only recently become available, this paper is limited to a study of the ECLS-K sample through third grade.

[2] The number of students sampled per classroom varied because of school sector (private schools with 12 or fewer kindergarters were eligible, while public schools with 24 or more kindergarteners were eligible), the need to oversample Asian Pacific Islanders, the inclusion of a twin subsample, and parental non-response. In general, the target number of children sampled at any one school (not including the second twin) was 24.

[3] We also performed estimation where we limited the sample to classes with at least five sampled students and obtained qualitatively similar results.

[4] NCES cautions against the estimation of absolute change in test scores because of the possibility that the metrics at different areas of the test score distribution are not comparable. This possibility provides an additional justification for our focus on percentile scores. Furthermore, we measure academic development on a relative scale because social development is measured on a relative scale. We have also estimated our models using IRT scores as dependent variables and obtained similar results.

[5] Because the underlying items used by NCES to construct its five scales of social development are proprietary, we were not able to perform our factor analysis on the underlying items themselves, which certainly would have been preferable from a scientific perspective.

[6] We also estimated these models separately for each adjacent pair of surveys. The results are very consistent across each pair.

REFERENCES

Alexander, Karl L., Doris R. Entwisle, and Maxine S. Thompson. 1987. "School Performance, Status Relations, and the Structure of Sentiment: Bringing the Teacher Back In." *American Sociological Review* 52: 665-682.

Armour, David T. 1976. *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* Santa Monica, CA: RAND Corporation.

Bidwell, Charles E., and John Kasarda. 1980. "Conceptualizing and Measuring the Effects of School and Schooling." *American Journal of Education* 8: 401-430.

Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science* 6: 42-57.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41: 778-820.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Oopportunity*. Washington, DC: U.S. Government Printing Office.

Downey, Douglas B., and Shana Pribesh. 2004. "When Race Matters: Teachers' Evaluations of Students' Classroom Behavior." *Sociology of Education* 77: 67-282.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation* 28: 465-498.

Ferguson, Ronald F. 1998. "Can Schools Narrow the Black-White Test Score Gap?" Pp. 218-374 in *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution.

Ferguson, Ronald F., and Helen F. Ladd. 1996. "How and Why Money Matters: An Analysis of Alabama Schools." Pp. 265-298 in *Holding Schools Accountable: Performance-based Reform in Education*, edited by Helen F. Ladd. Washington, DC: Brookings Institute.

Greenwald, Rob, Larry V. Hedges, and Richard D. Laine. 1996. "The Effect of School Resources on Student Achievement." *Review of Educational Research* 66: 361-396.

Hanushek, Eric. A. 1992. "The Trade-off Between Child Quantity and Quality." *Journal of Political Economy* 100: 84-117.

Hauser, Robert. M., William H. Sewell, and Duane F. Alwin. 1976. "High School Effects on Achievement." Pp. 309-341 in *Schooling and Achievement in American Society*, edited by William H. Sewell, Robert M. Hauser, and David L. Featherman. New York: Academic Press.

Hawkins, J. David, Jie Guo, Karl G. Hill, Sara Battin-Pearson, and Robert D. Abbott. 2001. "Long-Term Effects of the Seattle Social Development Intervention on School Bonding Trajectories." *Applied Developmental Science* 5: 225-236.

Hawkins, J. David, Rich Kosterman, Richard F. Catalano, Karl G. Hill, and Robert D. Abbott. 2005. "Promoting Positive Adult Functioning Through Social Development Intervention in Childhood." *Archives of Pediatrics and Adolescent Medicine* 159: 5-31.

Jencks, Christopher. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America.* New York: Harper and Row.

MacDonald, Kevin. 1985. "Early Experience, Relative Plasticity, and Social Development." *Developmental Review* 5: 99-121.

Milesi, Carolina, and Adam Gamoran 2006. "Effects of Class Size and Instruction on Kindergarten Achievement." *Educational Evaluation and Policy Analysis* 28: 287-313.

Murnane, Richard J. 1983. "Quantitative Studies of Effective Schools: What Have We Learned?" Pp. 193-209 in *School Finance and School Improvement: Linkages for the 1980s*, edited by Allan Odden and L. Dean Webb. Cambridge: Ballinger.

National Center for Educational Statistics. 1999. *Users' Manual for ECLS-K Base Year Data Files and Electronic Codebook*. Washington, DC: United States Department of Education.

Nelson, Charles A. 1999. "Neural Plasticity and Human Development." *Current Directions in Psychological Science* 8: 42-45.

Nye, Barbara, Spiros Konstantopoulos, and Larry V. Hedges. 2004. "How Large are Teacher Effects?" *Educational Evaluation and Policy Analysis* 26: 237-257.

Odden, Allan, Geoffrey Borman, and Mark Fermanich. 2004. "Assessing Teacher, Classroom, and School Effects, Including Fiscal Effects." *Peabody Journal of Education* 79: 4-32.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73: 417-458.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review, Papers and Proceedings*. 94: 247-252.

Rowan, Brian, Fang-Shen Chiang, and Robert J. Miller. 1997. "Using Research on Employees' Performance to Study the Effects of Teachers on Students' Achievement." *Sociology of Education* 70: 256-284.

Rowan, Brian, Richard Correnti, and Robert J. Miller. 2002. "What Large-Scale, Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools." *Teachers College Record* 104: 1525-1567.

Stiles, Joan. 2000. "Neural Plasticity and Cognitive Development." *Developmental Neuropsychology* 18: 237-272.

Yoshikawa, Hirokazu. 1995. "Long-Term Effects of Early Childhood Programs on Social Outcomes and Delinquency." *The Future of Children* 5: 51-75.

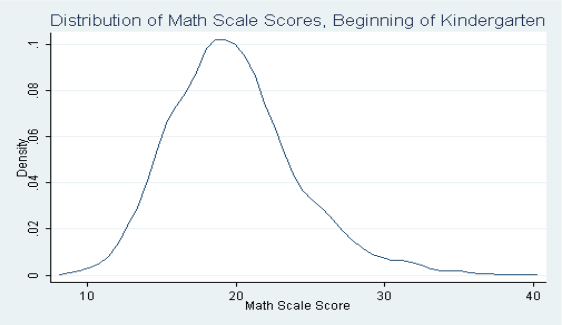
Table 1. Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Dependent Variables					
Δ externalizing behaviors, beginning of K-end of K	5613	.659	18.751	-72	72
Δ externalizing behaviors, beginning of K-end of 1	5613	.644	24.462	-72	74
Δ internalizing behaviors, beginning of K-end of K	5613	.170	24.461	-75	79
Δ internalizing behaviors, beginning of K-end of 1	5613	-.106	32.190	-75	82
Δ interpersonal/learning behaviors, beginning of K-end of K	5613	-.283	23.457	-88	91
Δ interpersonal/learning behaviors, beginning of K-end of 1	5613	-.367	30.409	-95	91
Δ interpersonal/learning behaviors, beginning of K hat-end of K	5613	-3.458	26.609	-81.490	75.321
Δ reading, beginning of K-end of K	5613	.197	20.042	-78	91
Δ reading, beginning of K-end of 1	5613	1.008	24.199	-79	82
Δ math, beginning of K-end of K	5613	.079	18.173	-82	78
Δ math, beginning of K-end of 1	5613	.743	22.156	-87	87
Δ reading, beginning of K-end of 3rd	4792	-.610	27.887	-93	86
Δ math, beginning of K-end of 3rd	4814	-.891	24.442	-78	91
Student Characteristics					
African-American	5613	0.144	0.351	0	1
Hispanic	5613	0.092	0.290	0	1
Asian	5613	0.026	0.158	0	1
Female	5613	0.501	0.500	0	1
SES	5613	49.172	28.743	1	100
Single parent family	5613	0.224	0.417	0	1
Age in months	5613	68.544	4.039	53.37	86.23
Biological mother present	5613	0.941	0.235	0	1
Only child	5613	0.155	0.362	0	1
Home language not English	5613	0.037	0.188	0	1
Student has a disability	5613	0.148	0.355	0	1
AFDC receipt	5613	0.088	0.283	0	1
Full day kindergarten	5613	0.580	0.494	0	1
Student retained in K	5613	0.030	0.168	0	1
Same teacher for K and 1	5613	0.026	0.150	0	1
Days between K academic assessments	5613	174.925	21.133	119	261
Days between K social assessments	5613	184.810	56.271	0	362
Missing days between K social assessments (Coded as 0 above)	5613	.080	0.271	0	1
Teacher Characteristics					
Between 35-49 years old	1036	.448	0.498	0	1
More than 50 years old	1036	.219	0.414	0	1
Novice teacher	1036	.034	0.180	0	1
1 year experience	1036	.046	0.210	0	1
2-5 years experience	1036	.175	0.380	0	1
Masters degree	1036	.335	0.472	0	1
Highest certification	1036	.641	0.480	0	1
Black	1036	.064	0.246	0	1

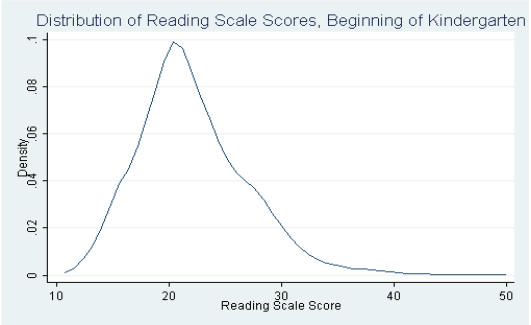
Hispanic	1036	.016	0.127	0	1
Instructional Styles					
Whole language scale	1036	16.808	5.156	0	24
Phonics scale	1036	20.878	3.589	0	24
Teaching math for understanding scale	1036	11.915	3.531	0	18
Drill-based math	1036	8.712	3.819	0	18

Source: Early Childhood Longitudinal Study – Kindergarten Cohort of 1999. See text for sample restrictions.

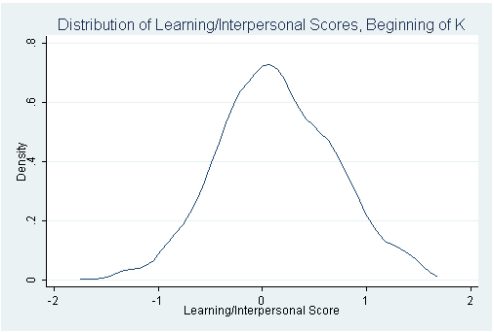
Figure 1. Distributions of Academic and Social Skills



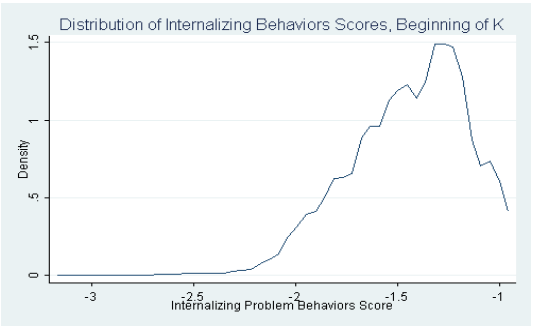
(A)



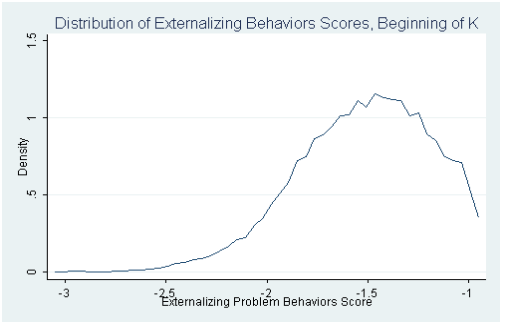
(B)



(C)



(D)



(E)

Figure 2. Conceptual Model

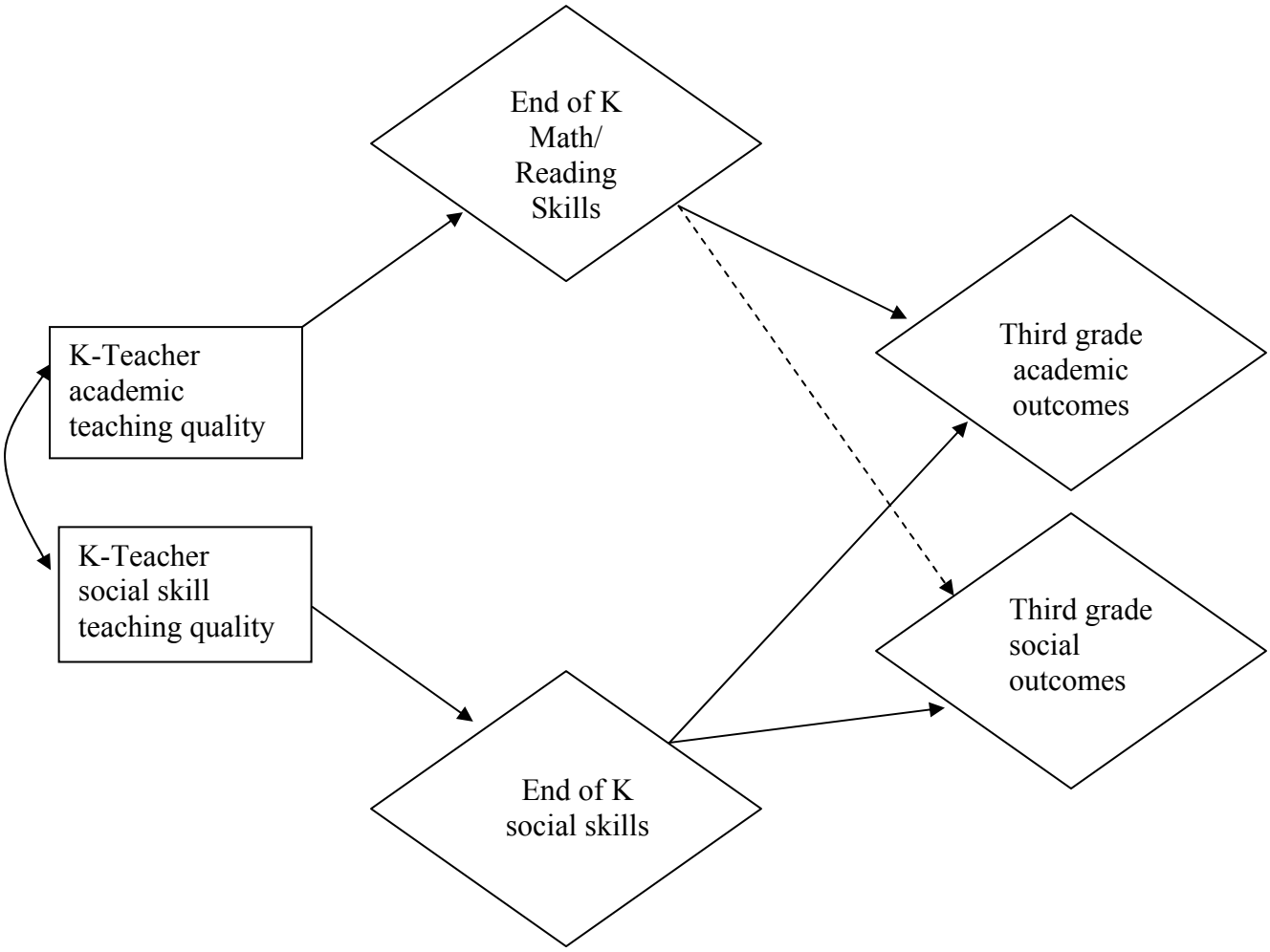


Table 2a. Growth Models Estimating Linear Effects of Social Development on Academic Development, Kindergarten Through Third Grade

	Panel A: Math			Panel B: Reading		
Variable	Learning-Interpersonal	Externalizing Behaviors	Internalizing Behaviors	Learning-Interpersonal	Externalizing Behaviors	Internalizing Behaviors
Score (t-1)	0.720*** (0.006)	0.732*** (0.006)	0.731*** (0.006)	0.667*** (0.006)	0.676*** (0.006)	0.677*** (0.006)
African-American	-4.524*** (0.482)	-4.482*** (0.484)	-4.718*** (0.484)	-3.779*** (0.522)	-3.792*** (0.526)	-4.032*** (0.526)
Hispanic	-0.265 (0.518)	-0.291 (0.520)	-0.283 (0.520)	-0.676 (0.562)	-0.759 (0.566)	-0.740 (0.566)
Asian	0.179 (0.637)	0.219 (0.639)	0.183 (0.640)	-0.409 (0.693)	-0.444 (0.698)	-0.443 (0.699)
Female	-3.411*** (0.298)	-2.947*** (0.298)	-2.697*** (0.291)	0.710* (0.324)	1.017** (0.327)	1.355*** (0.320)
SES	0.082*** (0.006)	0.085*** (0.006)	0.085*** (0.006)	0.105*** (0.006)	0.108*** (0.006)	0.109*** (0.006)
Single parent family	-0.331 (0.396)	-0.539 (0.398)	-0.620 (0.397)	-0.964* (0.432)	-1.122* (0.436)	-1.270** (0.435)
Learning/Interpersonal (t-1)	0.064*** (0.005)	—	—	0.062*** (0.006)	—	—
Externalizing (t-1)	—	0.033*** (0.006)	—	—	0.039*** (0.007)	—
Internalizing (t-1)	—	—	0.029*** (0.006)	—	—	0.023*** (0.006)
Age	0.004 (0.010)	0.006 (0.010)	0.006 (0.010)	-0.005 (0.012)	-0.003 (0.011)	-0.003 (0.012)
Age squared	-0.004*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.002** (0.001)
Intercept	10.082*** (0.512)	10.962*** (0.528)	11.164*** (0.519)	9.210*** (0.549)	9.941*** (0.568)	10.553*** (0.560)
n	4814			4792		

^ p≤.10; * p≤.05; ** p≤.01; *** p≤.001

Note: Standard errors are in parentheses. Score (t-1) is the lagged value of the appropriate dependent variable for each panel.

Table 2b. Growth Models Estimating Non-Linear Effects of Social Development on Academic Development, Kindergarten Through Third Grade

Variable	Panel A: Math			Panel B: Reading		
	(1)	(2)	(3)	(1)	(2)	(3)
Score (t-1)	0.720*** (0.006)	0.719*** (0.006)	0.719*** (0.006)	0.669*** (0.006)	0.669*** (0.006)	0.669*** (0.006)
African-American	-4.558*** (0.482)	-4.602*** (0.483)	-4.603*** (0.483)	-3.798*** (0.522)	-3.832*** (0.523)	-3.837*** (0.523)
Hispanic	-0.267 (0.519)	-0.266 (0.519)	-0.266 (0.519)	-0.643 (0.562)	-0.656 (0.562)	-0.667 (0.562)
Asian	0.180 (0.638)	0.182 (0.638)	0.183 (0.638)	-0.371 (0.693)	-0.396 (0.693)	-0.416 (0.693)
Female	-3.377*** (0.297)	-3.334*** (0.299)	-3.334*** (0.299)	0.785* (0.324)	0.789* (0.325)	0.789* (0.325)
SES	0.083*** (0.006)	0.083*** (0.006)	0.083*** (0.006)	0.105*** (0.006)	0.105*** (0.006)	0.105*** (0.006)
Single parent family	-0.340 (0.397)	-0.359 (0.398)	-0.357 (0.398)	-1.019* (0.432)	-0.996* (0.433)	-1.011* (0.433)
Age	0.005 (0.010)	0.003 (0.010)	0.004 (0.010)	-0.005 (0.012)	-0.006 (0.012)	-0.006 (0.012)
Age squared	-0.004*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.002** (0.001)
Quartile 1, Learning/Interpersonal	-4.698*** (0.437)	-5.064*** (0.522)	-5.035*** (0.533)	-4.192*** (0.475)	-4.080*** (0.569)	-4.222*** (0.580)
Quartile 2, Learning/Interpersonal	-2.663*** (0.418)	-2.797*** (0.435)	-2.816*** (0.441)	-3.137*** (0.464)	-3.073*** (0.483)	-2.974*** (0.490)
Quartile 3, Learning/Interpersonal	-0.733^ (0.414)	-0.764^ (0.416)	-0.772^ (0.417)	-1.191** (0.456)	-1.153* (0.459)	-1.111* (0.460)
Quartile 1, Externalizing Problems	—	0.681^ (0.399)	0.716^ (0.420)	—	0.062 (0.439)	-0.126 (0.464)
Quartile 1, Internalizing Problems	—	-0.256 (0.342)	-0.209 (0.384)	—	-0.443 (0.377)	-0.695 (0.425)
Quartile 1 in All Three Social Skills	—	—	-0.188 (0.705)	—	—	0.985 (0.775)
Intercept	15.353*** (0.596)	15.364*** (0.600)	15.357*** (0.601)	14.427*** (0.635)	14.496*** (0.639)	14.537*** (0.640)
n	4814			4792		

^ p≤.10; * p≤.05; ** p≤.01; *** p≤.001

Note: Standard errors are in parentheses. Score (t-1) is the lagged value of the appropriate dependent variable for each panel.

Table 3a. Intraclass Correlations for School and Teacher Effects: Random Effects Models

	Math		Reading		Learning/Interpersonal		Externalizing		Internalizing	
No controls	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers
Method 1	.187	.027	.219	.051	.037	.227	.041	.125	.042	.268
Method 2	.185	.014	.188	.021	.091	.014	.062	.006	.071	.030
With Controls	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers	Between schools	Between teachers
Method 1	.071	.017	.128	.041	.002	.264	.014	.111	.024	.196
Method 2	.107	.003	.118	.009	.084	.084	.067	.051	.074	.034
Method 3	.107	.003	.118	.009	.090	.019	.058	.012	.075	.025

Note: Control variables include the score or rating at the beginning of kindergarten, race, gender, socioeconomic status, family structure, the child's age, the presence of the biological mother, whether a language besides English is spoken at home, student disability, AFDC receipt, and full-day kindergarten. For Method 1, we include the number of days between assessments. For Methods 2 and 3, where scores or ratings at the end of first grade are the dependent variables, we also include student retention in kindergarten and whether the student has the same teacher again.

Table 3b. Teacher and School Effect Sizes: Fixed Effects Models

	Unadjusted Δ R			Adjusted Δ R		
	School Fixed Effects	Teacher Fixed Effects	Teacher Fixed Effects Controlling for School Fixed Effects	School Fixed Effects	Teacher Fixed Effects	Teacher Fixed Effects Controlling for School Fixed Effects
Math Method 1	0.224	0.304	0.205	0.162	0.184	0.087
Math Method 2	0.278	0.370	0.244	0.211	0.238	0.111
Reading Method 1	0.281	0.359	0.224	0.229	0.255	0.112
Reading Method 2	0.321	0.416	0.265	0.255	0.286	0.130
Learning/Interpersonal Method 1	0.301	0.473	0.365	0.231	0.386	0.310
Learning/Interpersonal Method 2	0.361	0.502	0.350	0.277	0.361	0.231
Learning/Interpersonal Method 3	0.371	0.501	0.337	0.277	0.326	0.171

Note: Control variables include the score or rating at the beginning of kindergarten, race, gender, socioeconomic status, family structure, the child's age, the presence of the biological mother, whether a language besides English is spoken at home, student disability, AFDC receipt, and full-day kindergarten. For Method 1, we include the number of days between assessments. For Methods 2 and 3, where scores or ratings at the end of first grade are the dependent variables, we also include student retention in kindergarten and whether the student has the same teacher in 1st grade.

Table 4. Effect of Moving a Student from the 25th to the 75th Percentile in the Teacher Effects Distribution

	Random Effects			Fixed Effects		
	Math	Reading	Learning/ Interpersonal	Math	Reading	Learning/ Interpersonal
Method 1	4.997	7.760	19.691	3.334	4.292	11.880
Method 2	2.099	3.636	11.107	4.254	4.982	8.853
Method 3	2.099	3.636	5.283	4.254	4.982	6.553

Table 5. Correlations between Empirical Bayes Estimates of Teacher Effects on Academic and Social Skills

Method	r (Math,Reading)	r (Math, Learning/ Interpersonal)	r (Reading, Learning/ Interpersonal)
Method 1	.338	.020	.081
Method 2	.364	.131	.122
Method 3	.364	.118	.132

n=1093 teachers

Table 6. Regression of End of 1st Grade Math, Reading, and Social Outcomes on Teacher Characteristics (Measured Using Method 3, with School and Teacher Random Effects Including Measured Teacher Characteristics)

Teacher characteristics	Math	Reading	Social
Between 35-49 years old	-0.123 (0.806)	0.693 (0.874)	0.785 (1.101)
More than 50 years old	-0.845 (0.984)	1.360 (1.068)	-0.038 (1.344)
Novice	-1.744 (1.909)	-0.252 (2.069)	-2.403 (2.605)
1 year experience	0.638 (1.560)	-0.599 (1.688)	-4.687* (2.141)
2-5 years experience	1.301 (0.955)	-1.025 (1.035)	-0.559 (1.305)
Masters degree	0.671 (0.695)	-1.385^ (0.755)	-0.519 (0.946)
Highest certification	1.354^ (0.751)	-0.458 (0.816)	2.007* (1.022)
African-American	1.906 (2.088)	1.282 (2.254)	-0.497 (2.839)
Hispanic	-7.549* (3.008)	-4.877 (3.251)	-2.355 (4.092)
African-American teacher*African-American student	-2.247 (2.747)	-1.996 (2.956)	0.821 (3.712)
Hispanic teacher*Hispanic student	11.885* (5.426)	10.140^ (5.863)	6.415 (7.279)

^ p≤.10; * p≤.05; ** p≤.01; *** p≤.001

Note: Standard errors are in parentheses. Control variables also include the score or rating at the beginning of kindergarten, race, gender, socioeconomic status, family structure, the child's age, the presence of the biological mother, whether a language besides English is spoken at home, student disability, AFDC receipt, full-day kindergarten, student retention in kindergarten, and whether the student has the same teacher in 1st grade.

Table 7. OLS Regression of Empirical Bayes Estimates of Teacher Effects on Instructional Characteristics of Classrooms

	Method 1			Method 2			Method 3
	Math	Reading	Social	Math	Reading	Social	Social
Whole language	0.003	0.018*	0.103^	0.002	0.003	-0.012	-0.000
	(0.005)	(0.008)	(0.059)	(0.002)	(0.004)	(0.026)	(0.008)
Phonics	0.000	0.024*	0.061	-0.001	0.010^	0.011	0.004
	(0.007)	(0.012)	(0.083)	(0.003)	(0.005)	(0.037)	(0.011)
Understanding math	0.004	-0.023^	0.054	-0.008*	-0.013*	-0.077^	-0.016
	(0.008)	(0.013)	(0.094)	(0.004)	(0.006)	(0.042)	(0.012)
Drill math	0.012	0.009	-0.041	-0.003	0.002	0.052	0.019^
	(0.008)	(0.012)	(0.084)	(0.003)	(0.005)	(0.038)	(0.011)
Intercept	-0.202	-0.606**	-3.287*	0.111^	-0.106	0.464	-0.045
	(0.134)	(0.210)	(1.495)	(0.059)	(0.091)	(0.671)	(0.196)
R^2	.006	.013	.008	.010	.008	.007	.005

n=1036 teachers

Note: Standard errors are in parentheses. The dependent variables in these analyses are the empirical Bayes estimates of teacher effects, which control for the score or rating at the beginning of kindergarten, race, gender, socioeconomic status, family structure, the child's age, the presence of the biological mother, whether a language besides English is spoken at home, student disability, AFDC receipt, and full-day kindergarten. For Method 1, we include the number of days between assessments. For Methods 2 and 3, where scores or ratings at the end of first grade are the dependent variables, we also include student retention in kindergarten and whether the student has the same teacher in 1st grade. These empirical Bayes estimates are also net of teacher characteristics included in Table 6.

Appendix A

Table A1. Loadings from Factor Analysis of Social Skills

Variable	Beginning of kindergarten	End of kindergarten	End of 1st grade
Interpersonal Skills	0.866	0.877	0.882
Self-Control	0.843	0.857	0.854
Approaches to Learning	0.769	0.752	0.739

Table A2. OLS Regression Predicting logit(Kindergarten Teachers' Ratings of Students at the Beginning of Kindergarten)

	Learning/Interpersonal Skills
African-American	-0.149* (0.068)
Hispanic	-0.156* (0.078)
Asian	0.117 (0.103)
Female	0.586*** (0.042)
SES	0.006*** (0.001)
Single Parent Family	-0.286*** (0.060)
Age in months	0.033*** (0.005)
Biological mother present	0.278** (0.092)
Only child	-0.206*** (0.059)
Home language not English	-0.060 (0.111)
Student has a disability	-0.131* (0.061)
Parent rating of child behavior	-0.559*** (0.033)
Public school	0.281*** (0.062)
AFDC Receipt	-0.244** (0.082)
Full day kindergarten	-0.067 (0.044)
Intercept	-3.150*** (0.384)
R^2	.153

^ p≤.10; * p≤.05; ** p≤.01; *** p≤.001

Note: Standard errors are in parentheses.

Table A3. Regression of End of 1st Grade Math, Reading, and Social Outcomes Using Method 3, with School and Teacher Random Effects But No Measured Teacher Characteristics

Method 3	Math	Reading	Social
Math, Beginning of K	0.649*** (0.011)	—	—
Reading, Beginning of K	—	0.603*** (0.011)	—
Predicted Social, Beginning of K	—	—	0.569*** (0.042)
African-American	-6.285*** (0.958)	-1.884 (1.023)	-2.961* (1.281)
Hispanic	-1.132 (1.016)	-0.105 (1.082)	1.488 (1.370)
Asian	-3.602** (1.338)	4.104** (1.427)	0.229 (1.799)
Female	-3.248*** (0.515)	0.983 (0.552)	3.812*** (0.962)
SES	0.081*** (0.011)	0.064*** (0.012)	0.038* (0.016)
Single Parent Family	0.868 (0.743)	-1.473 (0.790)	-0.979 (1.035)
Age in months	-0.025 (0.069)	-0.129 (0.072)	0.164 (0.097)
Biological mother present	0.140 (1.125)	0.176 (1.196)	-0.590 (1.544)
Only child	-2.676*** (0.729)	0.353 (0.776)	-0.328 (1.001)
Home language not English	3.387* (1.393)	2.195 (1.483)	3.206 (1.874)
Student has a disability	-1.873* (0.746)	-2.314** (0.792)	-1.346 (1.036)
AFDC Receipt	-1.207 (1.027)	-3.264** (1.093)	-2.529 (1.400)
Full day kindergarten	1.760* (0.792)	0.623 (0.870)	-0.291 (1.045)
Student retained in K	-13.863*** (1.601)	-19.505*** (1.705)	5.767** (2.149)
Student had same teacher in K and 1	-4.217* (1.889)	-2.803 (2.028)	-1.703 (2.538)
Intercept	18.006*** (4.920)	26.135*** (5.197)	8.044 (6.619)

^ p≤.10; * p≤.05; ** p≤.01; *** p≤.001

Table A4: Reliability of Academic and Social Measures

Academic Skills

Math	Beginning of K	End of K	End of 1st	End of 3rd
Beginning of K	1.0000			
End of K	0.8032	1.0000		
End of 1st	0.7096	0.7739	1.0000	
End of 3rd	0.6929	0.7389	0.7830	1.0000

n=4814

Reading	Beginning of K	End of K	End of 1st	End of 3rd
Beginning of K	1.0000			
End of K	0.7585	1.0000		
End of 1st	0.6574	0.7398	1.0000	
End of 3rd	0.5989	0.6174	0.6916	1.0000

n=4792

Teachers' Measures of Social Skills

Learning/ Interpersonal	Beginning of K	End of K	End of 1st	End of 3rd
Beginning of K	1.0000			
End of K	0.6756	1.0000		
End of 1st	0.4283	0.4560	1.0000	
End of 3rd	0.3883	0.4307	0.4924	1.0000

n=4142

Externalizing Problem Behaviors	Beginning of K	End of K	End of 1st	End of 3rd
Beginning of K	1.0000			
End of K	0.7209	1.0000		
End of 1st	0.5078	0.5479	1.0000	
End of 3rd	0.4648	0.5053	0.5350	1.0000

n=4170

Internalizing Problem Behaviors	Beginning of K	End of K	End of 1st	End of 3rd
Beginning of K	1.0000			
End of K	0.5545	1.0000		
End of 1st	0.2029	0.2476	1.0000	
End of 3rd	0.1996	0.2313	0.2451	1.0000

n=4170

Parents' Measures of Social Skills

Measure	Beginning of K/End of K	End of K/End of 1st	Beginning of K/End of 1st
Self-control	.597	.589	.552
Learning	.550	.544	.486
Social	.508	.484	.447
Impulsivity	.565	.556	.506
Sad/Lonely	.448	.427	.416

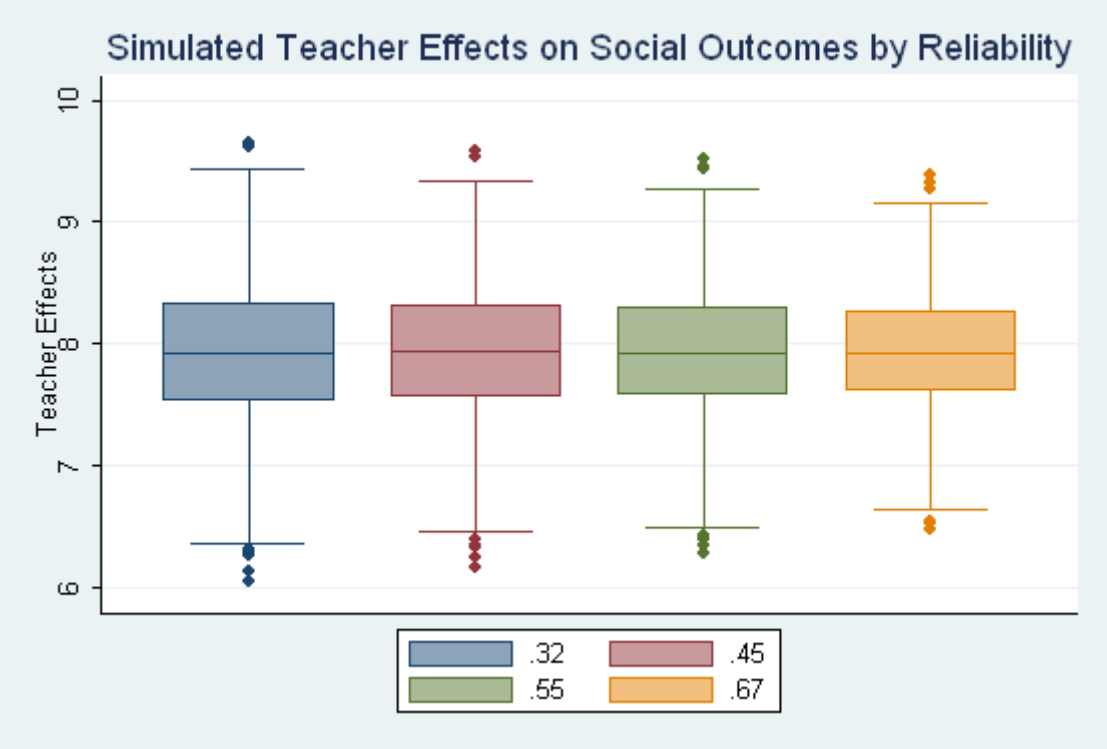
n=5,103

Appendix B: Reliability Simulation

The magnitude of estimated teacher effects on social development may be affected by the reliability of the social development measures in ECLS. To assess the extent to which reliability differences between the academic and social measures would alter our results, we generated a dataset with 7000 students and randomly assigned them to 700 teachers. Each student's observed score can be divided into three components: a "true score," an error component, and a teacher effect. We estimated the reliability of our social development measure by regressing each student's end of kindergarten score on their beginning of kindergarten score. In our data, the reliability of the social development measure is .68. From the distribution of true and error components implied by this reliability, each student was assigned a true score and an error score as well as a teacher effect. In this simulation, the social teacher effect was assigned a standard deviation of 8. We then decreased the reliability of the social measures to .55, .45, and .32 to determine how our estimates of teacher effects would be affected as a result, and ran each of these simulations 1,000 times.

Figure A1 below demonstrates that substantially decreasing the reliability of the social measures has only a small effect on our teacher effect estimates. We conclude that even large differences in the reliabilities of the academic and social measures would not alter our assertion that teacher effects on social development are at least as large as teacher effects on academic development.

Figure A1



ISERP
INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY

Institute for Social and Economic Research and Policy
Columbia University in the City of New York
420 West 118th Street
8th Floor, Mail Code 3355
New York, NY 10027
Tel: 212-854-3081
Fax: 212-854-8925
Email: iserp@columbia.edu
www.iserp.columbia.edu

DECEMBER 2007