

COGNITIVE AND NONCOGNITIVE PEER EFFECTS IN EARLY EDUCATION

Matthew Neidell and Jane Waldfogel

Abstract—We examine peer effects in early education by estimating value-added models with school fixed effects that control extensively for individual, family, peer, and teacher characteristics to account for the endogeneity of peer group formation. We find statistically significant and robust spillover effects from preschool on math and reading outcomes, but statistically insignificant effects on various behavioral and social outcomes. We also find that peer externalizing problems, which most likely capture classroom disturbance, hinder cognitive outcomes. Our estimates imply that ignoring spillover effects significantly understates the social returns to preschool.

I. Introduction

PEER effects have long been of interest in economics and other social sciences because they are a nonmarket interaction with far-reaching implications for school and community policies, such as school choice, ability tracking, desegregation, and antipoverty programs. Since peer effects may manifest themselves in various contexts, empirical research has focused on a wide range of outcomes—including academic performance, mental health, criminal activity, and use of public services—and ages—including primary and secondary school, higher education, and beyond.¹ Recent estimates of peer effects that use randomized and natural experiments to address the endogeneity of peer group formation generally find empirical support for peer effects, although estimates vary considerably depending on the outcome and age group studied.

In this paper, we examine peer effects in early education by looking at the effect of peer enrollment in preschool on children's outcomes in kindergarten and the early elementary grades. We focus on preschool because studies consistently demonstrate large private returns from early education on numerous cognitive outcomes.² Children may directly share the skills developed in preschool through social interactions in kindergarten, generating knowledge spillovers. Moreover, improved school readiness through early education may contribute to the pace of classroom learning in kindergarten, so the entire class indirectly benefits from peer enrollment in preschool. These direct and indirect peer effects may be particularly important at this

early age when environmental factors are so vital to development.

Since the peer effect we examine stems from a particular form of education (preschool), rather than inherent characteristics of peers, we interpret it as a spillover effect from early education. Previous research on spillover effects from education has typically focused on social returns from higher education, but the difficulty in finding plausibly exogenous variation in peer educational attainment has led to mixed empirical evidence.³ We contend that the variation in preschool enrollment used in this analysis is credible for identifying spillover effects. As growing amounts of state and federal funds are spent on public preschool programs, such as Head Start and public prekindergarten programs, knowledge of spillover effects is essential for assessing the efficient allocation of early education.

The perhaps more novel contribution of our study is the focus on noncognitive outcomes in addition to cognitive outcomes. Although preschool has demonstrated positive cognitive benefits, researchers often find negative social and behavioral consequences of preschool (Belsky et al., 2007). Children with limited self-control or discipline may unintentionally spread this behavior to peers, resulting in negative externalities. Given the growing evidence suggesting the importance of noncognitive skills in human capital acquisition and earnings (Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006), we also explore spillover effects from preschool enrollment on various social and behavioral outcomes.

Furthermore, children with behavioral problems may disturb teacher progress and hinder the learning of their classmates, so that peer behavior may affect academic performance. For example, an unruly child may require the teacher to focus more class time on discipline rather than on the dissemination of knowledge. Peer behavior is often the main rationale behind smaller class sizes (Lazear, 2001), and several studies of peer effects on academic performance recognize the role of peer disturbance (Hoxby, 2000; Hanushek et al., 2003; Ding & Lehrer, 2007), but with little direct empirical testing.⁴ We directly examine the spillover effects of broadly defined social and behavioral skills on academic achievement.

To account for the endogeneity of peers (Manski, 1993), we adopt a fixed effects, value-added approach, comparable to that employed by Hanushek et al. (2003), Arcidiacono and Nicholson (2005), and Ding and Lehrer (2007), using

Received for publication October 26, 2007. Revision accepted for publication October 8, 2008.

* Neidell: Columbia University and NBER; Waldfogel: Columbia University.

We thank Janet Currie, Sherry Glied, Cecilia Rouse, Greg Duncan, Steve Pischke, Michael Greenstone, two anonymous referees, and seminar participants at Columbia University and Princeton University for helpful comments, and Reina Kato for excellent research assistance. We both gratefully acknowledge funding from the Spencer Foundation, and J. W. from the John D. and Catherine T. MacArthur Foundation and Russell Sage Foundation.

¹ We describe previous research on peer effects in more detail in section II.

² See Currie (2001), Karoly et al. (2006), Cunha et al. (2006), Blau and Currie (2006), and Ludwig and Phillips (2007) for an overview.

³ See Acemoglu and Angrist (1999) and Moretti (2004) for a review of empirical evidence.

⁴ One notable exception is Figlio (2007), who focuses on peer disturbance in sixth-grade students in one school district in Florida.

data from the Early Childhood Longitudinal Study-Kindergarten Class. School fixed effects account for sorting into school districts by comparing children from classes within the same school. To account for selection into classes within a school, we estimate a value-added specification that controls for fall kindergarten scores, which are measured shortly after kindergarten begins and unlikely to be affected by peers. Moreover, a considerable advantage of focusing on kindergarten outcomes is the limited scope for student tracking, a potential source of bias in value-added models, because children are new to their environment. To account for common shocks during the kindergarten school year, we control extensively for individual, family, peer, and teacher characteristics.

We provide several pieces of evidence to support our empirical strategy. Peer preschool enrollment has small, statistically insignificant effects on fall-K scores, suggesting children are not assigned to peers based on initial ability. All available measures of teacher characteristics are uncorrelated with class average test scores, suggesting kindergarten students are not systematically assigned to teachers. Parents do not appear to compensate for their child's peer group by increasing other investments during the same time period. Once we include either school fixed effects or fall-K scores or both, our results are extremely robust to the inclusion of the numerous covariates available for our analysis, suggesting omitted variable bias is unlikely to plague our analysis. Finally, as a falsification test for value-added models (Rothstein, 2008), peer preschool composition in first grade is not significantly related to test score gains in kindergarten.

We find large, statistically significant effects of peer enrollment in preschool on math and reading outcomes. Peer preschool enrollment increases math and reading scores by 0.08 of a standard deviation in kindergarten, which implies the benefits from current preschool enrollment rates are understated by 16% to 25% if ignoring spillover effects. Furthermore, the impacts of kindergarten peer preschool enrollment persist in magnitude through first and third grades for math scores, though they decline somewhat for reading scores. These estimates imply that ignoring spillover effects significantly understates the social returns to preschool, supporting one rationale for public investment in preschool.

For noncognitive outcomes, preschool does not appear to directly spillover to peers' behaviors. Although we cannot rule out the possibility of preschool peer effects, we find imprecisely estimated spillover effects from peer preschool enrollment on all of our noncognitive outcomes. In terms of the impact of peer noncognitive outcomes on individual cognitive outcomes, we find that the mean of class behaviors does not have a statistically significant impact. However, the 75th and 90th percentile of peer externalizing problems, a noncognitive measure more likely to capture classroom disturbance, has an impact on math and reading scores, suggesting that only a handful of unruly students

may be sufficient for disrupting classroom learning. Despite these negative impacts, the spillover effects from preschool are positive on net.

II. Background Information

Preschool programs consist of a wide array of provisions, including community, school, state, and federal programs. An extensive body of research documents significant private returns to early enrichment programs, such as the High/Scope Perry Preschool Project and Carolina Abecedarian, and Head Start, a public preschool program for disadvantaged children. A large body of research focusing on general preschool programs (reviewed in Smolensky & Gootman, 2003, and Waldfogel, 2006) generally shows positive effects on cognitive outcomes, though they vary in magnitude across the types of program, but more mixed evidence on noncognitive outcomes, including both negative (Belsky et al., 2007; Magnuson, Ruhm, & Waldfogel, 2007) and positive effects (Puma et al., 2005). Despite the range of the results, the negative behavioral effects attracted major headlines and fueled debates over nonparental care (Carey, 2007).

Prior evidence of early education spillover effects is scant.⁵ Garces, Thomas, and Currie (2002) provide indirect evidence that spillover effects from Head Start may exist by comparing the effect of attending Head Start for older versus younger siblings. They posit that spillover effects are likely to flow from older to younger siblings because the older sibling is more likely to influence the younger and because learned parenting skills may benefit younger children more than older. They find no evidence of spillover effects on educational attainment or earnings when the children are adults, but some evidence on criminal activity.

Evidence on peer effects is vast, so we focus on recent studies most relevant to our study, recognizing that we omit several important studies. Since we examine peer effects during kindergarten, we first focus on studies that examined peer effects of primary school-aged children. Hanushek et al. (2003) and Hoxby (2000) found sizable endogenous peer effects on math scores in grades 3 to 6 in Texas. Lefgren (2004) found small but statistically significant endogenous peer effects on reading scores of third and sixth graders in Chicago public schools. In terms of exogenous peer effects, Angrist and Lang (2004) found insignificant peer effects from minority peers in Boston on math, reading, and language scores in the third, fifth, and seventh grades. Ammermueller and Pischke (2006) found positive peer effects of books owned on fourth-grade reading scores across several countries in Europe. These studies suggest that peer effects at these young ages exist, but they depend on the context and specific type of peer effect examined.

⁵ Estimates of spillover effects from higher education are pervasive (see Moretti, 2004, for a review), but differing approaches for dealing with the endogeneity of schooling have yielded different results.

Also relevant to our analysis are peer effect studies that focus on noncognitive outcomes. While most focus on crime, some using the Moving to Opportunity (MTO) experiment focus on mental health as well. Several studies using the MTO found neighborhood effects significantly reduced violent crimes by teens (Ludwig, Duncan, & Hirschfeld, 2001) and improved behavior problems for boys (Katz, Kling, & Liebman, 2001), though these effects disappeared or changed directions for boys and improved for girls in later follow-ups (Kling, Ludwig, & Katz, 2004, 2007; Sanbonmatsu et al., 2006). Bayer, Hjalmarsson, and Pozen (2007) found strong peer effects on criminal activity based on the composition of prison mates. Like cognitive outcomes, these peer effects differ by context, but they suggest the possibility of effects fading over time. Our study aims to add to this rich literature by examining peer effects on very young children, where effects may be particularly important, and by simultaneously examining cognitive and noncognitive outcomes.

III. Data

We use the Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), a nationally representative sample of kindergartners followed through first and third grades (National Center for Education Statistics, 2004). The ECLS–K contains detailed information collected through direct child assessments, parent interviews, and teacher and school questionnaires. We use the K–3 longitudinal file, which contains 17,401 children, and weight all analyses to account for survey nonresponse so our estimates are representative of the kindergarten class.⁶

A. Outcome Variables

For cognitive outcomes, we use standardized scores on mathematics and reading assessments administered at each survey wave. These assessments were designed to measure the age-specific achievement of the child, with scores standardized to have a mean of 50 and standard deviation of 10.

Given the difficulty in reliably measuring behavioral and social-emotional outcomes, we use several measures available from both teachers and parents. Teachers rated each individual student on four behavioral and social-emotional skills: self-control (controlling temper, respecting others' property, accepting peer ideas, and handling peer pressure), interpersonal skills (getting along with people; forming and maintaining friendships; helping other children; showing sensitivity to the feelings of others; and expressing feelings, ideas, and opinions in positive ways), externalizing problems (frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities), and internalizing problems (presence of anxiety, loneliness, low

self-esteem, and sadness). Externalizing problems are most consistent with the notion of peer disturbance (Lazear, 2001), so we pay particular attention to this measure.

Teachers' ratings of individual children may be subjectively reported relative to the average behavior of the class. For example, a generally disruptive child may be rated favorably in a class with numerous unruly peers but unfavorably in a class with few unruly peers. Therefore, we also use the parent's rating of his or her child's self-control as an outcome measure. Although the parent's rating may reflect the child's behavior at home rather than at school, the value-added specification limits this concern as long as the change in behavior at home from fall to spring correlates well with the change in behavior at school over the same time period.

These measures are adapted from the Social Skills Rating Scale, a widely used survey technique for detecting social and behavioral problems in the classroom. Each construct averages a series of questions rated on a scale of 1 (never) to 4 (very often), so a high score on self-control and interpersonal skills reflects a favorable outcome, and a high score on externalizing or internalizing problems reflects an unfavorable outcome. Their use in predicting future non-cognitive functioning is subject to debate, but these scales have high construct validity as assessed by test-retest reliability, internal consistency, interrater reliability, and correlations with more advanced behavioral constructs (Elliott et al., 1988) and are considered the most comprehensive social skill assessment that can be widely administered in large surveys such as the ECLS–K (Demaray et al., 1995).

B. Preschool Enrollment

Based on responses to the type of care arrangements in the year before kindergarten and the number of hours spent in each, the ECLS–K created a composite variable indicating the primary type of care. We use any center-based care (Head Start, day care, nursery school, preschool, and pre-kindergarten) to define preschool enrollment, leaving parental care and relative or nonrelative care as the alternative. Although each type of care may have different effects on children, the sample size prevents us from isolating peer effects from each source, so we examine the average peer effect across all types of care.

C. Control Variables

The ECLS–K contains detailed data on the family environment. Although we do not always measure specific inputs into human capital production and the full history of investments, we observe numerous proxies that reflect time and money available for families to invest at specific survey waves. For the mother, we include information on her current employment status, employment status at the period surrounding birth, educational attainment, immigration status, and age at which she gave birth. For the child, we

⁶ Results are generally unaffected by the use of sampling weights.

include gender, race/ethnicity, and birthweight. For household characteristics, we use income, number of siblings, presence of father in the household, father's employment status and education, central city or suburban residence, the number of grandparents the child has a close relationship with, whether English is spoken at home, and if anyone received food stamps or WIC. For direct measures of inputs, we include the number of books owned, number of records, tapes, or CDs, and whether the parent reads to the child every day.

D. Teacher and Class Variables

To account for teacher and classroom quality, we use data from the fall-K teacher questionnaire, which contains data on the class size and the teacher's background. The background measures include age, education, experience teaching both kindergarten and any other grade, years of tenure at the current school, gender, race, and type of teaching certificate. Furthermore, teachers were asked whether they enjoy teaching, would choose teaching again, and think that they make a difference in children's lives by their teaching. We created the variable "love teaching" equal to 1 if they responded yes to all three questions.⁷ While there are likely unmeasured components of teaching quality, these variables are widely accepted measures of teacher quality (Rivkin, Hanushek, & Kain, 2005), and our results are highly insensitive to their inclusion.

E. Peer Characteristics

Studies typically use population measures of peer characteristics, but such measures are unavailable at the class level. Instead, we compute class average enrollment in preschool and all other peer characteristics directly from the ECLS-K using the teacher identification number to identify students within the same class. Based on the number of students linked to each teacher and teacher-reported class size, the ECLS-K sampled 41% of students within a class on average.⁸ Although this greatly increases the quantity of peer characteristics we can control for, it induces measurement error that may bias our results. Fortunately, we can adjust our coefficient estimates for the undersampling of students (Ammermueller & Pischke, 2006), described in more detail below.

For computing effect sizes, the standard deviations of the peer variables are also measured with error. We unfortunately cannot adjust the standard deviations of these peer variables because they require two reports for preschool attendance (Ammermueller & Pischke, 2006). This pre-

cludes us from computing effect sizes from a 1 standard deviation change in peer preschool enrollment, a common metric used to compare results across studies. We can, however, use the standard deviation of 10 for math and reading scores to compute effect sizes from certain changes in preschool enrollment because this is a nationally normed test (normed scores are unavailable for the noncognitive outcomes).

F. Sample Characteristics

Starting from a possible sample of 17,401 children, we exclude children who repeated kindergarten (665), with preschool data missing (2,897), with only 1 child per class (so peer averages could not be computed) (416), and with test scores, teacher identification, and school identification unavailable (480), leaving a possible sample of 12,943 children. Of the covariates, a large number of cases were missing for all teacher and class variables and some parental responses (maternal age at birth, immigration status, grandparents close, early maternal employment, and number of books owned) (3,278), so we impute them using single imputation by chained equations to preserve sample size (Van Buuren et al., 2006).⁹ For the nonimputed covariates, we use casewise deletion, leaving a final sample of 12,501 with math scores, 11,840 with reading scores, and sample sizes ranging from 11,446 to 12,070 for the noncognitive outcomes.¹⁰

IV. Empirical Strategy

A. Econometric Model

To estimate peer effects, we focus on the following regression equation:

$$y_{icd}^s = \beta_0 pre_{icd} + \beta_1 \overline{pre_{(-i)cd}} + \beta_2 x_{icd} + \beta_3 \overline{x_{(-i)cd}} + \beta_4 z_{icd} + \beta_5 y_{icd}^f + \alpha_d + \eta_{cd} + \varepsilon_{icd}, \quad (1)$$

where y is the child's cognitive or noncognitive outcome in kindergarten, s indicates spring, f indicates fall, i is the individual, c is the classroom, and d is the school. pre indicates whether the child was enrolled in preschool and $\overline{pre_{(-i)cd}}$ is the mean enrollment of the class (not including the index child). x are individual and family-level characteristics, and z are classroom and teacher-specific characteristics. α_d is a school fixed effect. The error term consists

⁹ We choose not to impute the dependent variables or preschool variable because they do not appear to be missing at random, a necessary assumption for imputation to yield unbiased estimates. Furthermore, because we do not impute preschool, we perform single rather than multiple imputation, as standard errors for the preschool variables remain valid (Allison, 2002).

¹⁰ Children excluded from the analysis typically have lower test scores, worse noncognitive measures, and lower preschool enrollment, and they come from more disadvantaged backgrounds (not shown), suggesting our results may not generalize to the entire kindergarten cohort. Without further assumptions about the cause of missing observations, there is little we can do to overcome this concern.

⁷ These measures are from fall-K, so it is unlikely the teacher's responses have been affected by current students.

⁸ Our computed averages compare favorably with three teacher-reported averages available (51.0% versus 51.5% boys, 15.7% versus 17.6% African American, and 22.9% versus 19.2% Hispanic), although the latter two may differ from self-reports for reasons other than measurement error—teachers may misclassify the race and ethnicity of students.

of a group-specific component (η_{cd}) and an individual, idiosyncratic component (ε_{icd}). We cluster all standard errors at the classroom level. Our main hypothesis to test is that $\beta_1 = 0$.

We omit the class test score ($\overline{y_{(-i)cd}^f}$) from this equation to estimate the reduced-form peer effect of preschool: the direct effect on test scores plus the indirect effect through its impact on class test scores. We interpret this as the spillover effects from early education, a relevant parameter for understanding the efficient allocation of early education.

As previously mentioned, we use a sample of students in the class, rather than the entire class, to compute $\overline{pre_{(-i)cd}}$. Ammermueller and Pischke (2006) present a straightforward technique for adjusting the coefficient estimates to obtain consistent estimates free from measurement error; we present a version tailored to our model in the appendix. Consistency of peer effects estimate is given by

$$\hat{\beta}_{1adj} = \hat{\beta}_1 \left(\frac{\overline{N_{cd}} - 1}{\overline{n_{cd}} - 1} \right), \quad (2)$$

where n_{cd} is the average number of sampled students within a class (8.55) and N_{cd} is the average total number of students in the class (20.51), both observed in the data.¹¹

Endogeneity of β_1 arises if parents choose certain schools based on education preferences, so a school with higher-achieving students may reflect unobserved school quality or parental investments (Black, 1999). Moreover, children may sort into different classes based on the quality of the teacher or peers, so that the most able students end up with the best-quality teachers or peers. Alternatively, schools may assign children to different classrooms based on compatibility with the teacher or to obtain a mix of peers they deem optimal.¹²

Equation (1) has several features to aid in identification of β_1 . We include school fixed effects (α_d) to limit our comparison to children in different classrooms within the same school, which accounts for sorting into school districts. We also include fall kindergarten outcomes (y_{icd}^f) in equation (1), which are measured before children have had sufficient time to interact with their peers, so we examine how outcomes change over the course of the kindergarten year. Our estimated peer effect is therefore the correlation between the variation within schools in class preschool enrollment and the variation within schools of an individual's changes in test scores during kindergarten (conditional on the included covariates).

¹¹ Although class sizes vary in the sample, Ammermueller and Pischke (2006) demonstrate that using the overall means in their sample performs well for the level of variation in their sample, which is comparable to the level of variation we observe.

¹² Because we are focusing on exogenous peer effects and not endogenous peer effects, we do not focus on concerns regarding simultaneity bias.

B. Validity of Econometric Model

The summary statistics in table 1 highlight both the endogeneity of peer preschool enrollment and the strength of our methodology. Columns 1 and 2 show means and standard deviations for the individual-level variables and teacher characteristics for all children included in the analysis. Column 3 shows the difference in means of these variables for children in classes below and above the median class preschool enrollment of 59%, with p -values of the differences shown in column 4. Children with more peers in preschool have higher math and reading scores and generally more advantaged family backgrounds, such as higher parental income and education. Seventeen of the 28 covariates listed in the table have statistically significant differences for children in classes above and below the median, suggesting the importance of addressing the endogeneity of preschool enrollment.

Columns 5 and 6 highlight the strength of our empirical strategy by presenting these same differences and p -values after adjusting for school fixed effects. Immediately evident is that balance of covariates is achieved: only 1 of the 41 covariates has a statistically significant difference.¹³ Furthermore, the magnitude of the differences declines substantially. For example, the difference in family income falls from nearly \$12,000 without school fixed effects to just over \$700 with school fixed effects. When we further adjust for baseline test scores in addition to school fixed effects (not shown), the same balancing occurs. This supports the notion that peer preschool enrollment within schools is uncorrelated with unobservable factors that affect children's outcomes, a necessary condition for obtaining causal estimates of peer effects.

We probe possible selection effects into classes based on peer preschool enrollment by replacing the dependent variable in equation (1) with individual-level fall-K scores (and omit it from the right-hand side). If selection into classes is based on preschool enrollment, then peer preschool enrollment will have a significant effect on fall-K scores. Results from this specification, shown in columns 1 and 2 of table 2, indicate that fall-K scores are uncorrelated with peer preschool enrollment, with coefficient estimates small in magnitude and very imprecise. This suggests that within schools, students do not appear selectively assigned to peer preschool enrollment based on their initial ability. This result, in conjunction with the balancing of the adjusted covariates, suggests that school fixed effects may alone be sufficient for identifying causal peer effects. In fact, as we demonstrate below, the value-added specification has little

¹³ This does not include the preschool variables, which by definition are different. It is important to note that although school fixed effects reduce potential bias, there must still be sufficient variation in peer preschool enrollment after including school fixed effects. The overall standard deviation of peer preschool is 0.258, and the within-school variation is 0.213, suggesting ample variation for obtaining precise estimates.

TABLE 1.—SUMMARY STATISTICS

	1	2	3	4	5	6
	Mean	s.d.	Difference	p-Value	Difference (FE)	p-Value (FE)
<i>Dependent variables</i>						
Math spring K	51.080	9.667	1.854	0.000	0.328	0.019
Reading spring K	50.960	9.583	1.840	0.000	0.326	0.022
Self-control spring K (teacher rating)	3.196	0.621	-0.055	0.004	-0.026	0.035
Externalizing problems spring K	1.655	0.641	0.053	0.002	0.028	0.014
Interpersonal relations spring K	3.136	0.632	-0.019	0.317	-0.019	0.119
Internalizing problems spring K	1.556	0.508	0.001	0.963	0.010	0.326
Self-control spring K (parent rating)	2.881	0.498	0.035	0.002	-0.010	0.217
<i>Class and teacher characteristics</i>						
Class average preschool	0.580	0.258	0.416	0.000	0.194	0.000
Class size	20.651	4.381	-0.160	0.436	0.042	0.473
Teacher age	41.727	9.978	-0.036	0.940	0.217	0.449
Years tenure at school	9.344	7.850	-0.070	0.855	0.164	0.474
Years teach K	9.154	7.714	0.059	0.879	0.039	0.870
Total years teaching	12.948	9.218	-0.123	0.788	0.011	0.969
Teacher male	0.015	0.123	-0.009	0.136	0.001	0.725
Teacher white	0.862	0.345	0.002	0.916	-0.007	0.433
Temporary teaching certificate	0.084	0.277	0.007	0.578	0.007	0.409
Alternate teaching certificate	0.014	0.120	0.006	0.251	-0.001	0.853
Regular teaching certification	0.223	0.416	0.030	0.127	0.006	0.600
Highest teaching certification	0.656	0.475	-0.048	0.034	-0.012	0.360
Some graduate school	0.335	0.472	-0.071	0.001	-0.003	0.797
Master's or more	0.374	0.484	-0.005	0.845	0.000	0.993
<i>Background characteristics</i>						
Individual preschool	0.580	0.494	0.415	0.000	0.193	0.000
Maternal age at birth	23.886	5.406	1.207	0.000	0.065	0.339
Number of grandparents close relationship	2.215	1.242	0.070	0.014	0.027	0.136
Number of siblings	1.436	1.062	-0.078	0.001	0.000	0.991
Family income (\$)	52,167	44,329	11,648	0.000	731	0.171
Number of books owned	76.706	59.694	6.536	0.000	1.111	0.141
Number of records, tapes, CDs owned	15.417	17.705	1.695	0.000	0.445	0.076
Child age	6.207	0.342	0.002	0.832	0.003	0.563
Parent read to child every day	0.458	0.498	0.035	0.002	0.008	0.285
Mom employed full time	0.461	0.499	-0.003	0.799	-0.004	0.583
Mom employed part time	0.222	0.416	0.006	0.503	0.001	0.853
Dad employed full time	0.718	0.450	0.000	0.992	0.007	0.276
Dad employed part time	0.026	0.159	0.000	0.956	0.001	0.615
Central city residence	0.351	0.477	0.011	0.635	0.000	0.157
Suburban residence	0.423	0.494	0.112	0.000	0.000	0.901
Male	0.508	0.500	-0.005	0.584	-0.007	0.296
Black	0.144	0.351	0.079	0.000	0.004	0.355
Hispanic	0.167	0.373	-0.076	0.000	-0.014	0.003
Mother HS dropout	0.124	0.329	-0.054	0.000	-0.006	0.223
Mother HS graduate	0.634	0.482	-0.058	0.000	0.000	0.979
Father HS dropout	0.102	0.303	-0.054	0.000	0.000	0.930
Father HS graduate	0.471	0.499	-0.078	0.000	-0.007	0.310
Immigrant	0.172	0.377	-0.047	0.000	-0.009	0.094
Father present	0.784	0.412	-0.015	0.177	0.004	0.481
Early maternal employment	0.755	0.430	0.010	0.272	0.000	0.941
Low birthweight (<5.5 lbs.)	0.074	0.262	-0.004	0.496	-0.005	0.165
English at home	0.902	0.297	0.055	0.000	0.005	0.292
Mother or child WIC	0.451	0.498	-0.086	0.000	-0.003	0.680
Receive food stamps	0.172	0.377	-0.011	0.290	-0.001	0.865

Note: All values are weighted by sampling probability. "Difference" is the difference in means of the variables for children in classes above versus below the median class preschool enrollment of 59%. "p-value" is from *t*-test of variables below or above the median that cluster on class. FE adjusts variables for school fixed effect. Number of individuals = 12,501; number of schools = 901; number of classes = 2,436.

impact on our estimates once we include school fixed effects.

Even if peers are randomly assigned, families may respond by compensating for low peer quality by providing other investments in their child, leading to spurious estimates of peer effects (Moffitt, 2001).¹⁴ To explore whether

such compensatory behavior exists, we examine whether peer preschool enrollment affects two parental investments available in the ECLS-K: the number of books the child owns and whether the parent reads to the child every day. We estimate equation (1) by using the parental investments in place of *y*, though we use reports at the end of first grade

¹⁴ Parents may recognize their child's peer quality either directly by knowing the classmates or indirectly through their child's interim academic performance. Changing peer groups by switching classes is un-

likely to arise—less than 3% of respondents in our sample change classes within a school in kindergarten for any reason.

TABLE 2.—TESTS OF PEER GROUP EXOGENEITY

	1 Fall-K Outcomes		3 Compensatory Behavior		5 Future Inputs	
	Math	Reading	Number of Books Owned	Read to Child Every Day	Math	Reading
Individual preschool	0.922 [0.166]***	0.96 [0.177]***	3.151 [2.602]	0.000 [0.010]	-0.095 [0.118]	-0.187 [0.128]
Class preschool	0.056 [0.318]	0.033 [0.357]	3.928 [4.002]	-0.010 [0.020]	-0.025 [0.222]	0.244 [0.255]
Number of individuals	12,501	11,840	11,481	11,526	10,463	9,923
Number of schools	901	891	900	900	886	874
Number of classes	2,436	2,275	2,450	2,450	2,349	2,194
Covariates						
School fixed effects	Y	Y	Y	Y	Y	Y
Fall-K score	N	N	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y
Group characteristics	Y	Y	Y	Y	Y	Y
Teacher and class characteristics	Y	Y	Y	Y	Y	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. Individual characteristics are listed in table 1; group characteristics are class means of individual characteristics (not including reference child); and teacher and class characteristics are listed in table 1. Columns 1 and 2 present estimates of equation (1) using fall-K test scores as the dependent variable. Columns 3 and 4 present estimates of equation (1) using measures of parental investments in place of test scores. Columns 5 and 6 present estimates of equation (1) measuring all covariates in the first grade. * significant at 10%; ** significant at 5%; *** significant at 1%.

instead of spring-K for the dependent variable, the earliest time they are asked again in the ECLS-K after the fall-K interview. Shown in columns 3 and 4 of table 2, both investment measures are not statistically significantly related to peer preschool enrollment. Although there are other ways parents can compensate for their children's environment, this evidence generally does not support that such behavior exists.

Although the value-added model aids in identifying peer effects, two specific assumptions must be met.¹⁵ First, the impact of prior inputs and endowments decays at a constant rate. Although we cannot test this assumption directly, we observe several proxies for historical inputs and endowment and demonstrate that our results are insensitive to their

inclusion, suggesting the decay rate assumption is unlikely to affect our estimates.

The second assumption is that contemporaneous omitted inputs are not correlated with previous test scores, which may arise if teachers are strategically assigned to students based on their previous performance. We do not think this poses a considerable threat to our model because kindergarten children are mostly in a new school for the first time where limited historical information about the student is available. Also, we use fall kindergarten scores as the baseline score, which is obtained shortly after students have been assigned to teachers, so school administrators may not be aware of children's ability when assigning them to teachers.

We also provide empirical support for this assumption in panel A of table 3, which presents results from a regression

¹⁵ See Hanushek et al. (2003), Todd and Wolpin (2003), and Rothstein (2008) for a full derivation of this model.

TABLE 3.—TEST OF TEACHER ASSIGNMENT: EFFECTS OF TEACHER CHARACTERISTICS ON CLASS AVERAGE FALL-K SCORES

	1 Math	2 Reading	3 Self-Control (Teacher)	4 Interpersonal Skills	5 Externalizing Problems	6 Internalizing Problems	7 Self-Control (Parent)
<i>A. Regress fall-K class average outcome on K teacher and class characteristics</i>							
Dependent variable: Mean of class scores							
<i>F</i> -test teacher and class characteristics = 0	1.00	1.17	0.78	1.29	0.38	0.73	0.54
Prob > <i>F</i>	0.45	0.29	0.69	0.20	0.98	0.75	0.91
Dependent variable: Variance of class scores							
<i>F</i> -test teacher and class characteristics = 0	0.33	1.36	0.64	0.96	0.44	0.71	1.12
Prob > <i>F</i>	0.99	0.17	0.83	0.49	0.96	0.77	0.34
Number of classes	2,413	2,252	2,331	2,331	2,386	2,356	2,441
<i>B. Regress spring-K class average outcome on first-grade teacher and class characteristics</i>							
Dependent variable: Mean of class scores							
<i>F</i> -test teacher and class characteristics = 0	1.95	1.97	1.22	1.89	1.18	1.61	1.02
Prob > <i>F</i>	0.03	0.02	0.26	0.03	0.29	0.08	0.43
Dependent variable: Variance of class scores							
<i>F</i> -test teacher and class characteristics = 0	0.47	0.83	0.86	0.76	0.89	1.42	0.52
Prob > <i>F</i>	0.93	0.62	0.59	0.69	0.55	0.15	0.91
Number of classes	2,634	2,516	2,720	2,716	2,731	2,712	2,685

Note: All regressions weighted by sampling probability, include teacher and class characteristics as described in table 1, and include school fixed effects. *F*-test has 9 degrees of freedom. * significant at 10%; ** significant at 5%; *** significant at 1%.

of the class average scores in fall-K on teacher and class characteristics in kindergarten and school fixed effects. Within each panel, the dependent variable is either the mean or the variance of the individual scores within a class; teachers may have classes with comparable mean levels of ability, but a more effective teacher may be assigned to students with a wider range of abilities. Based on an F -test of joint significance, these characteristics are uncorrelated with all class-level outcomes, suggesting that contemporaneous inputs from teachers are unlikely to bias our results.

Alternatively, if we perform this same analysis for later grades, we find evidence that teachers are systematically assigned to students. In panel B, we regress class average scores from spring-K on first-grade teacher and class characteristics and school fixed effects. Teacher characteristics are correlated with mean math and reading scores even within schools, and are correlated with some of the noncognitive outcomes. These results support recent findings that nonrandom teacher assignment poses a concern for value added models (Rothstein, 2008), though we contend only in grades after kindergarten.

As a general specification check of the value-added model, we regress current outcomes on future inputs. Future inputs should not be correlated with current test scores, so any evidence to the contrary suggests our model is misspecified (Rothstein, 2008). In columns 5 and 6 of table 2, we present estimates from a regression of spring kindergarten outcomes on fall kindergarten outcomes and first-grade peer and teacher characteristics (omitting kindergarten peer and teacher characteristics). The coefficient on first-grade peer preschool is not statistically distinguishable from 0, further supporting the validity of our empirical strategy.

As further support for this assumption, we exploit the richness of background information available in the ECLS to control for numerous inputs during the kindergarten year. We include in x various individual-level factors, and because we compute peer characteristics from these individual characteristics, we also control extensively for composition of peers in $\bar{x}_{(-i)cd}$. Furthermore, we include the teacher and classroom characteristics in z_{cd} . As we demonstrate below, our estimates of β_1 are extremely robust to the inclusion of these variables.

We note that peer preschool enrollment may be correlated with individual preschool enrollment, suggesting it may be difficult to distinguish whether the effects of peer preschool enrollment are due to own enrollment or enrollment of peers. Although individual enrollment may be endogenous, our value-added specification captures the immediate private returns to preschool in the fall-K score. As with other historical inputs, as long as there is constant decay in the private returns to preschool, our estimates for class preschool enrollment will reflect peer effects. We test this assumption by adding individual preschool to our regression and (1) testing whether the coefficient on individual preschool differs from 0 and (2)

assessing whether the coefficient on peer preschool changes. For cognitive outcomes, we find both that individual preschool is not statistically significant and including it does not change the estimated peer preschool effect. For noncognitive outcomes, we find the coefficient on individual preschool is statistically significant, but our estimates of the peer effects are unchanged by including it.¹⁶

We also note that the coefficient on individual preschool reflects the additional effects of individual preschool above and beyond any immediate effects, so it provides a partial assessment of the private returns to preschool. Unfortunately, the value-added approach precludes us from identifying the full private returns to preschool enrollment since the immediate returns are captured by the lagged test score.

V. Results

A. Cognitive Spillover Effects from Preschool

For both cognitive outcomes, we present results from four sets of models: without both school fixed effects (FEs) and fall-K scores, without school fixed effects but with fall-K scores, with school fixed effects but without fall-K scores, and with both school fixed effects and fall-K score (our preferred specification). Within each model, we estimate four specifications: the first includes only individual and peer preschool enrollment as covariates (pre_{icd} and $pre_{(-i)cd}$), the second adds individual-level covariates (x_{icd}), the third adds group-level covariates ($\bar{x}_{(-i)cd}$), and the fourth adds teacher and class characteristics (z_{cd}). We make several comparisons across specifications to underscore the strength of our empirical strategy.

Columns 1 to 4 of table 4 show results for math scores without adjustment for measurement error in the peer variables. Focusing on estimates without school fixed effects and fall-K scores (panel A), we find a statistically significant effect of peer enrollment in preschool of 3.3, which implies that going from no peers in preschool to all peers in preschool raises math scores by 3.3 points. This estimate, however, gets successively smaller in magnitude as we control for more covariates. In column 4, which includes all covariates, the estimate falls to 0.92, though it remains statistically significant at conventional levels. This changing pattern across specification suggests the likelihood of omitted variable bias, so we cannot make strong claims that estimates from column 4 are free from bias.

When we add fall-K scores but continue to omit school fixed effects (panel B), we find much less variability in estimates across the specifications: they range from 0.55 to 0.70, and all are statistically significant. The stability of estimates suggests that the value-added approach controls

¹⁶ Results are not shown, but are available from authors on request.

TABLE 4.—SPILLOVER EFFECTS OF PRESCHOOL ON SPRING KINDERGARTEN COGNITIVE OUTCOMES

	1	2	3	4	5	6	7	8
	Math				Reading			
<i>A. No school FE, no fall-K score</i>								
Individual preschool	1.958 [0.197]***	0.854 [0.167]***	0.847 [0.166]***	0.816 [0.165]***	1.662 [0.202]***	0.736 [0.182]***	0.705 [0.180]***	0.654 [0.178]***
Class preschool	3.300 [0.456]***	1.211 [0.327]***	0.992 [0.325]***	0.920 [0.324]***	2.791 [0.455]***	1.289 [0.378]***	1.069 [0.378]***	0.946 [0.372]**
R^2	0.02	0.30	0.31	0.31	0.02	0.22	0.23	0.24
<i>B. No school FE, fall-K score</i>								
Individual preschool	0.013 [0.111]	0.051 [0.112]	0.074 [0.112]	0.063 [0.112]	-0.162 [0.124]	-0.114 [0.126]	-0.102 [0.126]	-0.123 [0.125]
Class preschool	0.547 [0.236]**	0.612 [0.234]***	0.695 [0.241]***	0.663 [0.241]***	0.494 [0.265]*	0.631 [0.263]**	0.747 [0.274]***	0.679 [0.274]**
R^2	0.68	0.69	0.69	0.70	0.62	0.63	0.64	0.64
<i>C. School FE, no fall-K score</i>								
Individual preschool	1.207 [0.179]***	0.724 [0.165]***	0.719 [0.165]***	0.698 [0.164]***	1.081 [0.180]***	0.578 [0.170]***	0.557 [0.169]***	0.525 [0.169]***
Class preschool	0.751 [0.359]**	0.652 [0.312]**	0.588 [0.312]*	0.549 [0.311]*	0.841 [0.389]**	0.752 [0.343]**	0.597 [0.345]*	0.528 [0.344]
R^2	0.30	0.41	0.41	0.41	0.29	0.39	0.39	0.40
<i>D. School FE, fall-K score</i>								
Individual preschool	0.028 [0.106]	0.003 [0.106]	0.003 [0.106]	-0.002 [0.106]	-0.112 [0.112]	-0.149 [0.113]	-0.149 [0.113]	-0.163 [0.112]
Class preschool	0.524 [0.215]**	0.527 [0.213]**	0.517 [0.212]**	0.506 [0.213]**	0.608 [0.247]**	0.599 [0.244]**	0.544 [0.249]**	0.504 [0.248]**
R^2	0.74	0.75	0.75	0.75	0.71	0.72	0.72	0.72
Observations	Individuals = 12,501, schools = 901, classes = 2,436				Individuals = 11,840, schools = 891, classes = 2,275			
SUR joint test	$\chi^2(2) = 7.99$				$P > \chi^2 = 0.02$			
<i>Covariates</i>								
Individual characteristics	N	Y	Y	Y	N	Y	Y	Y
Group characteristics	N	N	Y	Y	N	N	Y	Y
Teacher and class characteristics	N	N	N	Y	N	N	N	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. See notes to table 2 for description of covariates. All columns contain regression results of equation (1) using spring-K math and reading scores as the dependent variable. Panels B and D include fall-K math and reading scores as independent variables. SUR joint test is a test of joint significance of class preschool on the math and reading equations from columns 4 and 8. * significant at 10%; ** significant at 5%; *** significant at 1%.

for many confounding factors. Additionally, the coefficient on individual preschool now becomes statistically insignificant, indicating that the value-added specification controls well for prior inputs and endowment, and most of the effect of preschool on cognitive outcomes is immediate.

When we add school FEs but omit fall-K scores (panel C), we also find estimates that range minimally (from 0.55 to 0.75) as we add more covariates and remain statistically significant at the 5% or 10% level. This suggests that once we account for sorting into school districts, there is little evidence of sorting into classes within schools based on preschool enrollment, as results from tables 1 through 3 supported. Importantly, despite an entirely different source of variation, these estimates are quite comparable to the value-added estimates without school FEs, providing further credibility to our estimation strategy.

In the value-added specification with school FEs (panel D), our estimates range even less as we add more controls, ranging from 0.53 to 0.51. The spillover estimate that includes only individual preschool, fall-K scores, and school FEs is 0.52, and including the full set of over 70 covariates slightly reduces the estimate to 0.51. When this is broken down further, our estimates change from 0.52 in

column 3 to 0.51 in column 4 when we add the measures of teacher quality and classroom characteristics, further supporting the notion that teacher sorting is unlikely to drive our results. The individual preschool effect now is essentially 0, suggesting our peer preschool variable is in fact capturing the impact of peers. The R^2 ranges from 0.74 to 0.75, suggesting we are controlling for a considerable amount of variability in math scores. Many of the control variables we add are highly predictive of children's achievement, so it is impressive that our spillover estimates are unaffected by their inclusion. In fact, comparing the estimates from column 4 of panel C to panel D, we see our R^2 increases from 0.41 to 0.75, but our estimates for peer preschool change from only 0.55 to 0.51. Despite the vast increase in explanatory power of the regression, the stability of our estimate further demonstrates the robustness of our methodology.

Turning to reading scores in columns 5 to 8, we find not only the same pattern of robustness across models and within specifications but estimates that are nearly identical in magnitude as the math scores. Estimates range from 0.50 to 0.61 in the specification with school fixed effects and fall-K scores and are statistically significant at the 5% level. This remarkable stability across important predictors and different sources of variation is

TABLE 5.—MAGNITUDE OF SPILLOVER EFFECTS

Δ preschool enrollment	(1)	(2)	<i>GG</i>	(3)	
	B_{1adj}/σ_y	$(B_{1adj}/\sigma_y) \times \Delta$ Preschool		Column 2/Private Returns <i>LP low</i>	<i>LP high</i>
0–100%	0.132	0.132	0.338	0.548	0.366
0–45.8%	0.132	0.076	0.155	0.252	0.167
45.8–100%	0.132	0.055	0.183	0.297	0.198

Notes: B_{1adj} is the measurement error adjustment given in equation (2) using coefficient estimates from column 4 from panel D of table 2. σ_y is the standard deviation of test score (=10).

supportive that our empirical strategy is uncovering causal spillover effects and not merely correlations.¹⁷

To assess the magnitude of these spillover effects, ideally we would like to compute effect sizes from a standard deviation change in peer preschool enrollment to facilitate comparisons to other peer effects studies. We can use equation (2) to adjust our coefficients for measurement error in peer preschool enrollment and use the normalized standard deviation of 10 for test scores, but unfortunately we cannot make an adjustment to the standard deviation of peer preschool enrollment with the data at hand.

As an alternative approach for assessing magnitude, we compute the spillover effects from various changes in preschool enrollment rates and compare them to estimated private returns from preschool programs; this yields estimates of the missed benefits from ignoring spillover effects. For example, we compute the spillover effects of moving from no children in preschool to the current rate of preschool enrollment, which gives estimates of the missed benefits from current preschool enrollment rates. For estimates of the private returns to preschool, we use estimates from Gormley and Gayer (2005) for private returns to prekindergarten and from Puma et al. (2005) for private returns to Head Start. Although our preschool variable captures a wide array of provisions and may not be directly comparable to these private returns,¹⁸ we choose these two because both have strong research designs and provide readily comparable effect sizes for cognitive outcomes. Gormley and Gayer (2005), who exploited the introduction of a universal prekindergarten program in Tulsa, Oklahoma, by comparing those born just after the cutoff date for

admission to those born just before, found cognitive outcomes of prekindergarten attendees improved by 0.39 of a standard deviation in kindergarten (denoted *GG* in table 5). The Head Start Impact Study (Puma et al., 2005), which randomly assigned children into Head Start, found statistically significant effect sizes ranging from 0.16 to 0.24 across several cognitive assessments measured at age 4. Since the estimates from Puma et al. (2005) reflect intent to treat, to maintain consistency with our estimates we instead use the treatment on the treated estimates from Ludwig and Phillips (2007) that range from 0.24 to 0.36 (denoted *LP low* and *LP high*, respectively, in table 5).

We compute the following spillover rates using estimates from our preferred specification: (1) going from 0% to 100% enrollment, (2) going from 0% to the current enrollment rate of 45.8%; and (3) going from the current enrollment rate to 100%.¹⁹ The third comparison is potentially important for understanding the impacts of universal preschool, although we recognize this makes the strong assumption that current quality of preschool remains unchanged under universal preschool. Table 5 shows computations for the three alternatives. Estimates from the second row imply that the benefits from current preschool enrollment rates are understated by 16% to 25% if ignoring spillover effects, and estimates from the third row imply that the benefits of moving from current preschool enrollment rates to universal preschool are understated by 18% to 30%. Results from this chart suggest that spillover effects are an important component of private returns; the social returns to preschool are considerably understated if spillover effects are overlooked.

B. Noncognitive Spillover Effects from Preschool

Turning to the behavioral and social-emotional outcomes, shown in table 6, we present results only from our preferred specification because we examine five different dependent variables, though we find similar robustness patterns across specifications as with the cognitive outcomes. In contrast to cognitive outcomes, we do not find statistically significant spillover effects on behavioral or social-emotional outcomes. In fact, for none of the five outcomes in any of the specifications do we find a statistically significant effect. We

¹⁷ We also estimate our model using the gain in test scores in kindergarten as the dependent variable rather than including fall-K scores as an independent variable. When we use our full specification with school fixed effects and all covariates, the estimate for math scores is 0.493 and for reading scores is 0.495; both are statistically significant at the 5% level. These results are also largely insensitive to excluding the covariates, ranging from 0.464 to 0.495 for math scores and 0.495 to 0.538 for reading scores. As an additional specification check, we also estimate models that include the class average fall-K scores (excluding the reference child) to control for unobserved group effects not accounted for in the group-level covariates. Estimates are largely unaffected by including the class average scores: the estimate for math scores changes slightly from 0.506 to 0.471 and on reading scores is virtually unchanged at 0.507.

¹⁸ We estimated models that include both Head Start (HS) and non-HS enrollment of class, and although this stretches our sample and we lose considerable precision, estimates are comparable for the two groups. For math, the coefficient on non-HS is .478 and for HS is .340 (compared to .5 overall), and for reading, it is .434 and .623, respectively (compared to .5 overall).

¹⁹ We cannot simply add the last two to get the first one because peer effects are multiplier effects. We recognize this assumes linear peer effects, which we explore in section V.E.

TABLE 6.—SPILLOVER EFFECTS OF PRESCHOOL ON SPRING KINDERGARTEN NONCOGNITIVE OUTCOMES

	1	2	3	4	5
	Self-Control (Teacher)	Interpersonal Skills	Externalizing Problems	Internalizing Problems	Self-Control (Parent)
Individual preschool	-0.042 [0.010]***	-0.043 [0.011]***	0.032 [0.010]***	0.019 [0.009]**	-0.007 [0.009]
Class preschool	-0.029 [0.024]	-0.038 [0.025]	0.014 [0.021]	0.028 [0.020]	-0.023 [0.018]
R^2	0.52	0.5	0.6	0.42	0.43
Number of individuals	11,636	11,446	11,919	11,759	12,070
Number of schools	881	883	889	887	901
Number of classes	2,324	2,321	2,380	2,342	2,462
SUR joint test	$\chi^2(5) = 5.17$	$P > \chi^2 = 0.39$			
Covariates					
School fixed effects	Y	Y	Y	Y	Y
Fall-K score	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y
Group characteristics	Y	Y	Y	Y	Y
Teacher and class characteristics	Y	Y	Y	Y	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. See notes to table 2 for a description of covariates. All columns contain regression results of equation (1) using spring-K noncognitive measures as the dependent variable. SUR joint test is a test of joint significance of class preschool on all noncognitive equations in columns 1 through 5. *significant at 10%; **significant at 5%; ***significant at 1%.

also perform a joint test across all five outcomes by performing a seemingly unrelated regression, and the p -value from this test is 0.39. Although the estimates are not statistically significant, the effect sizes of moving from no peers in preschool to all peers in preschool (using the sample standard deviations) are comparable in magnitude to the effect sizes for the cognitive outcomes, though we cannot be certain these effect sizes are correct because the standard deviations of the noncognitive outcomes are measured with error.

We also find of interest that the individual preschool effect persists despite including fall-K scores, suggesting that children who attended preschool have social or behavioral problems that increase over the course of the kindergarten year. Although the value-added approach does not yield estimates of the full behavioral effect—fall-K scores absorb initial behavioral effects—finding a continued effect is strongly suggestive that adverse individual-level behavioral effects from preschool exist.

C. Impact of Peer Noncognitive Development on Individual Cognitive Outcomes

We next examine whether peer noncognitive development has an effect on cognitive outcomes by including in equation (1) the class average from fall-K of each of the noncognitive measures. If peer behavior disrupts classroom learning, we expect to find positive coefficients on self-control and interpersonal skills and negative coefficients on externalizing and internalizing problems. We find from columns 1 and 4 of table 7 that the class averages of the behavioral and social-emotional outcomes have no statistically significant effect on cognitive outcomes, according to both individual t -tests and a joint F -test that circumvents concerns regarding multicollinearity of the noncognitive measures. Furthermore, including these measures has little effect on our estimates of the preschool spillover effects.

One concern with using the mean of class behaviors is that it may only take one disruptive student to inhibit classroom instruction (Lazear, 2001), and this may not be captured in the mean. We instead include the 75th and 90th percentiles of class behaviors, which will more reliably detect the presence of any disruptive children. In the subsequent columns in table 7 (columns 2 to 3 and 5 to 6), we find a statistically significant negative effect for externalizing problems on math scores, the behavioral measure most consistent with disruptive behaviors. For reading scores, none of the noncognitive outcomes are statistically significant individually, but a joint F -test suggests they have a statistically significant impact on reading scores. We find little difference between estimates from the 75th and 90th percentiles for both math and reading scores, which suggests that only a handful of unruly peers are sufficient for disrupting the academic progress of their classmates.

We also note that the coefficient on peer preschool becomes slightly larger in these specifications as well. This is not surprising because, as previously mentioned, table 4 produces reduced-form estimates of the impact of peer preschool enrollment by combining the direct effect of peer preschool enrollment plus the indirect effect through its impact on class averages. Therefore, since we have partialled out the negative indirect effect from class averages in table 7, the direct effect from peer preschool enrollment increases. This also indicates that although the results from table 7 suggest some negative spillover effects from preschool, the reduced-form estimates from table 4 indicate that the net spillover effects from preschool are positive.

D. Persistence of Spillover Effects

Even if students show initial gains from peers, these effects may disappear over time as children age, given evidence of fade-out effects from Head Start (Currie &

TABLE 7.—SPILLOVER EFFECTS OF NONCOGNITIVE OUTCOMES ON SPRING KINDERGARTEN COGNITIVE OUTCOMES

	1	2	3	4	5	6
		Math			Reading	
Class preschool	0.600 [0.239]**	0.679 [0.238]***	0.666 [0.238]***	0.482 [0.279]*	0.613 [0.276]**	0.599 [0.277]**
Class self-control	-0.089 [0.312]	0.198 [0.245]	0.095 [0.231]	0.265 [0.338]	0.185 [0.256]	0.246 [0.233]
Class interpersonal skills	-0.099 [0.259]	-0.057 [0.216]	0.054 [0.212]	-0.016 [0.290]	0.292 [0.248]	0.330 [0.235]
Class externalizing problems	-0.315 [0.232]	-0.526 [0.168]***	-0.555 [0.140]***	0.418 [0.275]	-0.177 [0.208]	-0.091 [0.163]
Class internalizing problems	-0.204 [0.230]	0.076 [0.164]	0.060 [0.144]	0.115 [0.277]	-0.115 [0.196]	-0.062 [0.155]
Mean or percentile of noncognitive outcomes	Mean	75th	90th	Mean	75th	90th
F-test noncognitive = 0	0.93	5.28	9.24	0.78	3.45	4.21
Prob > F	0.45	0.00	0.00	0.54	0.01	0.00
Number of individuals	11,257	11,257	11,257	10,768	10,768	10,768
Number of schools	885	885	885	871	871	871
Number of classes	2,256	2,256	2,256	2,134	2,134	2,134
<i>Covariates</i>						
School fixed effects	Y	Y	Y	Y	Y	Y
Fall-K score	Y	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y
Group characteristics	Y	Y	Y	Y	Y	Y
Teacher and class characteristics	Y	Y	Y	Y	Y	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. See notes to table 2 for description of covariates. All columns contain regression results of equation (1) using spring-K math and reading scores as the dependent variable. Self-control is from the teacher report. * Significant at 10%; ** significant at 5%; *** Significant at 1%.

Thomas, 1995). In table 8, we explore the persistence of effects by looking at the effects of peer preschool enrollment on spring first- and third-grade outcomes. For math scores, we find comparable point estimates (0.68 for first grade and 0.60 for third grade) to the spillover estimates in the spring of kindergarten. Estimates are less precise, which could be due in part to the smaller sample that results from attrition, though they remain statistically significant for the first-grade results. For reading scores, we find some evidence the impacts fade over time, with estimates falling to 0.36 in first grade and 0.37 in third grade, though this difference is not statistically significant. This slightly fading impact is consistent with the negative coefficient on indi-

vidual preschool (from table 4), which suggests that private returns to preschool on reading decreases over time. Both sets of estimates are also generally insensitive to the inclusion of control variables.²⁰

²⁰ For behavioral and social-emotional outcomes (not shown), the estimates remain statistically insignificant and become considerably smaller in magnitude for externalizing and internalizing problems. We also explore the heterogeneity of results by both estimating quantile regressions and interacting peer preschool with several covariates, but find few significant differences. For more details, see Neidell and Waldfogel (2008).

TABLE 8.—SPILLOVER EFFECTS OF PRESCHOOL ON COGNITIVE OUTCOMES IN FIRST AND THIRD GRADES

	1	2	3	4	5	6	7	8
		Math				Reading		
<i>First grade</i>								
Class preschool	0.660 [0.386]*	0.614 [0.369]*	0.657 [0.372]*	0.679 [0.370]*	0.490 [0.456]	0.473 [0.433]	0.368 [0.441]	0.361 [0.434]
Number of individuals	10,084	10,084	10,084	10,084	9,530	9,530	9,530	9,530
Number of schools	894	894	894	894	884	884	884	884
Number of classes	2,376	2,376	2,376	2,376	2,213	2,213	2,213	2,213
<i>Third grade</i>								
Class preschool	0.514 [0.428]	0.509 [0.403]	0.576 [0.388]	0.602 [0.391]	0.403 [0.502]	0.358 [0.476]	0.407 [0.464]	0.374 [0.461]
Number of individuals	10,062	10,062	10,062	10,062	9,466	9,466	9,466	9,466
Number of schools	894	894	894	894	883	883	883	883
Number of classes	2,375	2,375	2,375	2,375	2,211	2,211	2,211	2,211
<i>Covariates</i>								
School fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Fall-K score	Y	Y	Y	Y	Y	Y	Y	Y
Individual characteristics	N	Y	Y	Y	N	Y	Y	Y
Group characteristics	N	N	Y	Y	N	N	Y	Y
Teacher and class characteristics	N	N	N	Y	N	N	N	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. See notes to table 2 for a description of covariates. All columns contain regression results of equation (1) using spring first- or third-grade math and reading scores as the dependent variable. * significant at 10%. ** significant at 5%; *** significant at 1%.

TABLE 9.—NONLINEAR SPILLOVER EFFECTS OF PRESCHOOL ON SPRING KINDERGARTEN COGNITIVE OUTCOMES

	1 Math	2 Reading
Class preschool 25–50%	0.329 [0.188]*	–0.029 [0.208]
SUR test	$\chi^2(2) = 1.94$	$P > \chi^2 = 0.16$
Class preschool 50–75%	0.514 [0.203]**	0.152 [0.223]
SUR test	$\chi^2(2) = 1.73$	$P > \chi^2 = 0.19$
Class preschool 75–100%	0.476 [0.203]**	0.406 [0.237]*
SUR test	$\chi^2(2) = 0.06$	$P > \chi^2 = 0.80$
Number of individuals	12,501	11,840
Number of schools	901	891
Number of classes	2,436	2,275
<i>Covariates</i>		
School fixed effects	Y	Y
Fall-K score	Y	Y
Individual background characteristics	Y	Y
Group background characteristics	Y	Y
Teacher and class characteristics	Y	Y

Note: Robust standard errors that cluster on class in brackets. All regressions weighted by sampling probability. See notes to table 2 for a description of covariates. All columns contain regression results of equation (1) using spring-K math and reading scores as the dependent variable. Class preschool 0–25% is reference category. SUR test is a test that the class preschool indicators on the math equation from column 1 equal that of the reading equations from column (2). * significant at 10%; ** significant at 5%; *** significant at 1%.

E. Nonlinear Effects

Thus far we have assumed linear peer effects from peer preschool enrollment, so we now explore the possibility of nonlinear peer effects. Nonlinear effects may occur if, for example, peer effects do not arise until a critical mass of classmates attends preschool and additional preschool students beyond the critical mass provide little marginal benefit. Previous studies, summarized in Ammermueller and Pischke (2006), find varying evidence on nonlinearities, suggesting they may be specific to the dependent variable or peer characteristic explored.

In table 9, we explore nonlinear effects from peer preschool enrollment flexibly by specifying indicator variables for whether 25% to 50%, 50% to 75%, and 75% to 100% of the class was enrolled in preschool (with 0 to 25% the reference category). Peers have a statistically significant effect on math achievement when preschool enrollment reaches 25% to 50%, shown in column 1. The effect increases minimally beyond 50%, though it remains statistically significant, suggesting little marginal benefit from additional preschool students once a certain level is reached. Shown in column 2, for reading scores, however, the effect of peers is small and not statistically significant when 25% to 50% of students are enrolled in preschool, but increases monotonically until it becomes statistically significant when 75% to 100% of peers are enrolled. Although the differences in estimates for the same quartile across test subjects are not statistically significant when estimating the equations jointly, the disparity by subject is intriguing, though not particularly surprising, given previous research demonstrating that nonlinear effects appear context specific.

VI. Conclusion

This paper explores spillover effects from early education programs by examining the effect of peer enrollment in preschool on children's outcomes in kindergarten. To address the endogeneity of peer group formation, we estimate value-added models with school fixed effects that control extensively for individual, family, peer, and teacher characteristics. Although kindergarten outcomes are of interest in their own right, focusing on this age group strengthens our empirical strategy because of less sorting into classes based on prior outcomes. Numerous sensitivity analyses support the interpretation of our estimates as causal estimates of peer effects.

We find robust, significant spillover effects from preschool on math and reading scores that appear to persist through the third grade. We find little evidence to support direct social and behavioral spillovers, but we find that peer noncognitive development, particularly externalizing problems, affects student achievement, suggesting an indirect route through which preschool has negative impacts. The social returns to preschool, however, are positive on net.

These results have three implications. First, significant spillover effects from preschool suggest a potentially suboptimal allocation of preschool enrollment. Second, our results that peer externalizing behaviors have an impact on individual cognitive outcomes supports the key assumption through which smaller class sizes improve academic achievement (Lazear, 2001). Third, and most generally, our evidence of strong peer effects as early as kindergarten supports contentions that policies such as school choice and ability tracking will have significant consequences for the academic achievement of children.

REFERENCES

- Acemoglu, Daron, and Joshua Angrist, "How Large Are the Social Returns to Educations? Evidence from Compulsory Schooling Laws," NBER working paper no. 7444 (1999).
- Allison, Paul, *Missing Data: Series: Quantitative Applications in the Social Sciences* (Thousand Oaks, CA: Sage, 2002).
- Ammermueller, Andreas, and Jorn-Steffen Pischke, "Peer Effects in European Primary Schools: Evidence from PIRLS," NBER working paper no. 12180 (2006).
- Angrist, Joshua, and Kevin Lang, "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *American Economic Review* 94 (2004), 1613–1634.
- Arcidiacono, Peter, and Sean Nicholson, "Peer Effects in Medical School," *Journal of Public Economics* 89 (2005), 327–350.
- Bayer, Patrick, Randi Hjalmarsson, and David Pozen, "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections," NBER working paper no. 12932 (2007).
- Belsky, Jay, Deborah Lowe Vandell, Margaret Burchinal, K. Alison Clarke-Stewart, Kathleen McCartney, and Margaret Tresch Owen, "Are There Long-Term Effects of Early Child Care?" *Child Development* 78 (2007), 681–701.
- Black, Sandra, "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics* 114 (1999), 577–599.
- Blau, David, and Janet Currie, "Who's Minding the Kids? Preschool, Day Care, and After School Care," in E. Hanushek and F. Welch (Eds.), *The Handbook of Economics of Education* (New York: North Holland, 2006).

- Carey, Benedict, "Poorer Behavior Is Linked to Time in Day Care," *New York Times*, March 26, 2007.
- Cunha, Flavio, James Heckman, Lance Lochner, and Dimitryi Masterov, "Interpreting the Evidence on Life Cycle Skill Formation," in E. Hanushek and F. Welch (Eds.), *The Handbook of Economics of Education* (New York: North-Holland, 2006).
- Currie, Janet, "Early Childhood Intervention Programs: What Do We Know?" *Journal of Economic Perspectives* 15 (2001), 213–238.
- Currie, Janet, and Duncan Thomas, "Does Head Start Make a Difference?" *American Economic Review* 85 (1995), 341–364.
- Demaray, Michelle, Stacey Ruffalo, John Carlson, R. T. Busse, Amy Olson, Susan McManus, and Amy Leventhal, "Social Skills Assessment: A Comparative Evaluation of Six Published Rating Scales," *School Psychology Review* 24 (1995), 648–671.
- Ding, Weili, and Steven Lehrer, "Do Peers Affect Student Achievement in China's Secondary Schools?" this REVIEW 89 (2007), 300–312.
- Elliott, Stephen, Frank Gresham, Terry Freeman, and George McCloskey, "Teacher and Observed Rating of Children's Social Skills: Validation of the Social Skills Rating Scales," *Journal of Psychoeducation Assessment* 6 (1988), 152–161.
- Figlio, David, "Boys Named Sue: Disruptive Children and Their Peers," *Education Finance and Policy* 2 (2007), 376–394.
- Garces, Eliana, Duncan Thomas, and Janet Currie, "Longer Term Effects of Head Start," *American Economic Review* 92 (2002), 999–1012.
- Gornley, William, and Ted Gayer, "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program," *Journal of Human Resources* 40 (2005), 533–558.
- Hanushek, Eric, John Kain, Jacob Markman, and Steven Rivkin, "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (2003), 527–544.
- Heckman, James, and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review* 91 (2001), 145–149.
- Heckman, James, Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24 (2006), 411–482.
- Hoxby, Caroline, "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER working paper no. 7867 (2000).
- Karoly, Lynn, Peter Greenwood, Susan Everingham, Jill Hoube, M. Rebecca Kilburn, C. Peter Rydell, Matthew Sanders, and James Chiesa, *Investing in Our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Interventions* (Santa Monica, CA: RAND, 2006).
- Katz, Lawrence, Jeffrey Kling, and Jeffrey Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics* 116 (2001), 607–654.
- Kling, Jeffrey, Jens Ludwig, and Lawrence Katz, "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics* 120 (2005), 87–130.
- , "Experimental Analysis of Neighborhood Effects," *Econometrica* 75 (2007), 83–119.
- Lazear, Edward, "Educational Production," *Quarterly Journal of Economics* 116 (2001), 777–803.
- Lefgren, Lars, "Educational Peer Effects and the Chicago Public Schools," *Journal of Urban Economics* 56 (2004), 169–191.
- Ludwig, Jens, Greg Duncan, and Paul Hirschfield, "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment," *Quarterly Journal of Economics* 116 (2001), 655–679.
- Ludwig, Jens and Deborah Phillips, "The Benefits and Costs of Head Start," NBER working paper no. 12973 (2007).
- Magnuson, Katherine, Christopher Ruhm, and Jane Waldfogel, "Does Prekindergarten Improve School Preparation and Performance?" *Economics of Education Review* 26 (2007), 33–51.
- Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60 (1993), 531–542.
- Moffitt, Robert, "Policy Interventions, Low-Level Equilibrium, and Social Interactions," in S. Durlauf and P. Young (Eds.), *Social Dynamics* (Cambridge, MA: MIT Press, 2001).
- Moretti, Enrico, "Human Capital Externalities in Cities," in V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (New York: North-Holland, 2004).
- National Center for Education Statistics, *User's Manual for the ECLS-K Third Grade Public Use Data File and Electronic Code Book* (Washington, DC: U.S. Department of Education, 2004).
- Neidell, Matthew, and Jane Waldfogel, "Cognitive and Non-Cognitive Peer Effects in Early Education," NBER working paper no. 14277 (2008).
- Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, and Michael Lopez, "Head Start Impact Study: First Year Findings" (Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, 2005).
- Rivkin, Steven, Eric Hanushek, and John Kain, "Teachers, Schools, and Academic Achievement," *Econometrica* 73 (2005), 417–458.
- Rothstein, Jesse, "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," Princeton University mimeograph (2008).
- Sanbonmatsu, Lisa, Jeffrey Kling, Greg Duncan, and Jeanne Brooks-Gunn, "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment," *Journal of Human Resources* 41 (2006), 649–691.
- Smolensky, Eugene, and Jennifer Appleton Gootman, *Working Families and Growing Kids: Caring for Children and Adolescents* (Washington, DC: National Academies Press, 2003).
- Todd, Petra, and Kenneth Wolpin, "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* 113 (2003), F3–F33.
- Van Buuren, S., J. Brand, C. Groothuis-Oudshoorn, and D. Rubin, "Fully Conditional Specification in Multivariate Imputation," *Journal of Statistical Computation and Simulation* 76 (2006), 1049–1064.
- Waldfogel, Jane, *What Children Need* (Cambridge, MA: Harvard University Press, 2006).

APPENDIX

In this section we sketch the measurement error adjustment in equation (2) based on the derivation in Ammermueller and Pischke (2006).²¹ Suppose we are interested in estimating the following equation (omitting other covariates for clarity):

$$y_{icd} = \beta_1 w_{cd} + \varepsilon_{icd},$$

where $w_{cd} = \overline{pre}_{(-i)cd}$ is the average preschool enrollment less the index child for the entire class of students (N_{cd}). We observe a sample of students only in the class (n_{cd}) and denote the measured average preschool enrollment by w_{cd}^* . The estimate of β_1 converges to

$$p \lim \hat{\beta}_1 = \frac{\text{cov}(y, w^*)}{\text{var}(w^*)} = \frac{\text{cov}(\beta_1 w + \varepsilon, w^*)}{\text{var}(w^*)} = \frac{\beta_1 \text{cov}(w, w^*)}{\text{var}(w^*)}.$$

We must define the covariance and variance terms in the above equation, where P_{-i} is the set of all students in the class and S_{-i} is the set of sampled students in the class:

$$\begin{aligned} \text{var}(w^*) &= \text{var}\left(\sum_{s_{-i}} \frac{pre_i}{n_{cd} - 1}\right) = \left(\frac{1}{n_{cd} - 1}\right)^2 \text{var}\left(\sum_{s_{-i}} pre_i\right) \\ &= \left(\frac{1}{n_{cd} - 1}\right)^2 (n_{cd} - 1) \text{var}(pre_i) = \left(\frac{1}{n_{cd} - 1}\right) \text{var}(pre_i) \end{aligned}$$

²¹ The main difference is we assume no measurement error in the individual-level covariates and therefore derive results for a bivariate regression.

$$\begin{aligned}
\text{cov}(w, w^*) &= \text{cov}\left(\sum_{p-i} \frac{pre_i}{N_{cd}-1}, \sum_{s-i} \frac{pre_i}{n_{cd}-1}\right) = \left(\frac{1}{N_{cd}-1}\right) \\
&\times \left(\frac{1}{n_{cd}-1}\right) \sum_{p-i} \sum_{s-i} \text{cov}(pre_i, pre_i) \\
&= \left(\frac{1}{N_{cd}-1}\right) \left(\frac{1}{n_{cd}-1}\right) (n_{cd}-1) \text{var}(pre_i) \\
&= \left(\frac{1}{N_{cd}-1}\right) \text{var}(pre_i).
\end{aligned}$$

Plugging these terms in gives

$$p \lim \hat{\beta}_1 = \frac{\beta_1 \left(\frac{1}{N_{cd}-1}\right) \text{var}(pre_i)}{\left(\frac{1}{n_{cd}-1}\right) \text{var}(pre_i)} = \beta_1 \left(\frac{n_{cd}-1}{N_{cd}-1}\right).$$

We can use the measured averages for N_{cd} and n_{cd} from our sample to adjust our estimate of β_1 to obtain consistent estimates of the peer effect:

$$\hat{\beta}_{1adj} = \hat{\beta}_1 \left(\frac{\overline{N_{cd}}-1}{\overline{n_{cd}}-1}\right).$$