

IZA DP No. 9305

**What Differences a Day Can Make:
Quantile Regression Estimates of the Distribution of
Daily Learning Gains**

Michael S. Hayes
Seth Gershenson

August 2015

What Differences a Day Can Make: Quantile Regression Estimates of the Distribution of Daily Learning Gains

Michael S. Hayes
Rutgers University–Camden

Seth Gershenson
*American University
and IZA*

Discussion Paper No. 9305
August 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

What Differences a Day Can Make: Quantile Regression Estimates of the Distribution of Daily Learning Gains*

Recent research exploits a variety of natural experiments that create exogenous variation in annual school days to estimate the average effect of formal schooling on students' academic achievement. However, the extant literature's focus on average effects masks potentially important variation in the effect of formal schooling across the achievement distribution. We address this gap in the literature by estimating quantile regressions that exploit quasi-random variation in the number of school days between kindergarten students' fall and spring tests in the nationally representative Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K). The marginal effect of a typical 250-day school-year on kindergarten students' math and reading gains varies significantly, and monotonically, across the achievement distribution. For example, the marginal effect on the 10th percentile of the reading achievement distribution is 0.9 test score standard deviation (SD), while the marginal effect on the 90th percentile is 2.1 test score SD. We find analogous results for math achievement.

JEL Classification: I2

Keywords: education production function, school year length, quantile regression, ECLS-K

Corresponding author:

Seth Gershenson
School of Public Affairs
American University
4400 Massachusetts Avenue, NW
Washington DC, 20016-8070
USA
E-mail: gershens@american.edu

* The authors thank Dave Marcotte for providing helpful comments. Any remaining errors are our own.

1. Introduction

Instructional time is increasingly recognized as an important input in the education production function. Knowing the impact of instructional time on student achievement is critical to ensuring the efficient allocation of scarce resources and to conducting cost-benefit analyses of a variety of education policies regarding modifications to the school calendar, weather-related school closures, and student and teacher absences. Accordingly, recent research has employed a variety of quasi-experimental methods to examine the causal relationship between school days (or unexpected school closures, teacher absences, or student absences) and student achievement (Aucejo and Romano, 2013; Bellei, 2009; Clotfelter, Ladd, and Vigdor, 2009; Fitzpatrick et al., 2011; Goodman, 2014; Hansen, 2008; Herrmann and Rockoff, 2012; Leuven et al., 2010; Marcotte and Hemelt, 2008; Parinduri, 2014; Pischke, 2007; Sims, 2008). Generally, this literature finds that additional school days have a positive impact on student achievement. For example, Fitzpatrick et al. (2011) exploit quasi-random variation in test dates in the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) to identify the average effect of formal schooling on achievement gains and find that kindergarten reading scores increase by 1.6 test score standard deviations (SD) during a standard 250 day school year.

However, this literature focuses exclusively on the identification and estimation of average effects. While average effects are interesting and add to our understanding of the relationship between instructional time and student outcomes, they overlook potential variation across the achievement distribution in the relationship between instructional time and student achievement (Bitler, Domina, Penner, and Hoynes, 2015; Eide and Showalter, 1998). We begin to fill this gap in the literature by extending the identification strategy pioneered by Fitzpatrick et al. (2011) to the quantile regression context. Specifically, we use nationally representative,

student-level data on kindergarteners in the ECLS-K to estimate quantile regressions that relate instructional time to student achievement at various points in the achievement distribution. Consistent with Fitzpatrick et al. (2011), we find the average effect of a typical school-year period (250 days) on kindergarteners' reading achievement to be 1.6 test score SD. Our novel contribution, though, is providing evidence that such effects actually vary monotonically across the achievement distribution. For example, the marginal effect on the 10th percentile of the reading achievement distribution is 0.9 test score SD, the marginal effect on the 90th percentile is 2.1 test score SD, and these two effects are statistically significantly different from one another. Analyses of math achievement yield similar results. This suggests that high- and low-achieving students experience different gains from instructional time, a nuance that has been overlooked to date in this literature.

2. Data and Methods

We use data from the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K), which was administrated by the National Center for Education Statistics (NCES). The sample of more than 20,000 children from about 1,000 kindergarten programs (i.e., schools) was designed to be nationally representative of the cohort that entered kindergarten in the 1998-99 academic year.¹ All children were surveyed in both the fall and spring of kindergarten.

The ECLS-K administered age-appropriate reading and mathematics tests in both the fall and spring semesters.² Following Fitzpatrick et al. (2011), in all subsequent analyses test scores

¹ Certain subgroups of the population were oversampled. We follow Fitzpatrick et al. (2011) in reporting unweighted estimates. However, in Online Appendix A, we show that weighting the regressions by NCES-provided sampling weights that adjust for the ECLS-K's nonrandom sampling frame yields qualitatively similar results.

² See Fitzpatrick et al. (2011) and Quinn (2014) for further discussion of the tests.

are standardized to have mean zero and SD one using the mean and SD of the fall kindergarten test scores for all children in the sample.³ Importantly, in both the fall and spring semesters, ECLS-K assessments were administered to different students on different days. Differences in test dates across schools, across classrooms within schools, and even across students within classrooms are common in the data, as a relatively small number of ECLS-K administrators individually met with each student to perform the assessment. Fitzpatrick et al. (2011) show that test dates and time between ECLS-K tests were essentially random. Following Fitzpatrick et al. (2011), we leverage this exogenous variation in instructional time between tests to identify the distribution of quantile treatment effects of instructional time on achievement gains.

The analytic sample contains 16,050 kindergarteners who were first-time kindergarteners and who had both fall and spring test scores.⁴ Descriptive statistics for these students are reported in Table 1. The average gains between the fall and spring in reading and math were 1.18 and 1.13 SD, respectively. On average, tests were about 187 days apart. Again, as shown in Fitzpatrick et al. (2011), the identifying variation comes from the quasi-random variation in the ECLS-K test dates. Because “instructional time between tests” was as good as randomly assigned, the impact of instructional time on achievement gains can be estimated via straightforward linear and quantile regressions of the form

$$Y_i^{Spring} - Y_i^{Fall} = \alpha + \beta SchoolYears_i + u_i, \quad (1)$$

where Y represents student i 's standardized kindergarten math or reading tests and $SchoolYears$ represents the number of “typical 250-day school years” to which the student was exposed

³ Alternative scalings (e.g., unstandardized IRT scale scores) (Quinn, 2014) yield qualitatively similar results; see Online Appendix B.

⁴ Sample sizes are rounded to the nearest 50, as per NCEs rules for restricted-use ECLS-K data.

between test dates (Fitzpatrick et al., 2011).⁵ We also estimate augmented versions of equation (1) that control for the student and classroom characteristics described in Table 1. The results are robust to conditioning on these covariates, which provides additional evidence that the test dates are essentially random (Fitzpatrick et al., 2011).⁶

We estimate Equation (1) by OLS, essentially replicating Fitzpatrick et al. (2011), and then estimate corresponding quantile regressions that take the right hand side of equation (1) as the linear index. The baseline models are linear in *SchoolYears*, as Fitzpatrick et al. (2011) find average school-year learning rates in the ECLS-K to be approximately linear. Online Appendix E shows that learning rates are approximately linear across the achievement distribution as well. Standard errors are clustered at the school level and computed via 500 bootstrap replications.

3. Results

Table 2 reports baseline OLS and quantile regression estimates of the effect of a 250-day school year student achievement gains. The OLS estimates in columns 1 and 3 are similar to those reported in Fitzpatrick et al. (2011). Consistent with Fitzpatrick et al. (2011), we see no significant difference between the coefficient estimates in models that do and do not condition on

⁵ Equation (1) is known as a gain-score model in the value-added literature, which is the specification estimated in Fitzpatrick et al. (2011). Alternatively, equation (1) could be specified as a lag-score model, in which Y^{Fall} is included as a regressor on the right hand side. See Quinn (2014) for a thoughtful discussion of the different interpretations of the two models. However, as shown in Online Appendix C, lag-score versions of equation (1) yield qualitatively similar results. Following Fitzpatrick et al. (2011), *SchoolYears* is computed by dividing the number of days between the spring and fall tests by 250. The 250 number comes from assuming a 180 day school year and adding in weekends and holidays.

⁶ The main quantile treatment effect estimates are also robust to conditioning on school fixed effects (FE). Specifically, we implement the non-additive quantile FE estimator proposed by Powell (2015). These estimates are directly comparable to the baseline bivariate quantile regression estimates and are reported in Online Appendix D.

observed student and classroom characteristics. Again, this suggests that time between tests is essentially random.

The quantile regression estimates show that there are significant differences in the impact of instructional time on academic achievement at different points in the achievement distribution.⁷ Interestingly, the estimated effects monotonically increase with achievement levels. This is clearly illustrated in Figure 1, which plots the OLS estimates, each quantile estimate, and the point estimates' 95% confidence intervals for both reading and math achievement. Figure 1 also shows that quantile estimates above the 90th percentile and below the 15th percentile are significantly different from the OLS estimates.

It is also possible that the effect of instructional days varies across classroom settings or other student characteristics. Particularly relevant to the current context is the distinction between half-day and full-day kindergarten programs, as some research suggests that full-day programs produce larger learning gains (Cannon et al., 2006; DeCicca, 2007; Gibbs, 2014). Then we might expect to see that time in full-day kindergarten programs has a larger impact on achievement than time in half-day programs. We test this hypothesis by augmenting equation (1) to include *Half-day*×*SchoolYears* interaction terms, but find no significant differences in linear or quantile regressions. Similarly, and consistent with Fitzpatrick et al. (2011), we find no evidence of heterogeneous effects across any student characteristics.

4. Discussion

The literature on the impact of instructional time on student achievement focuses entirely on average effects estimated via linear models. The current study contributes to this literature by

⁷ Eide and Showalter (1998) find similar patterns in quantile regressions that do not explicitly address the possible endogeneity of school-year length.

estimating quantile regressions that examine whether the impact of instructional time on student achievement varies across the achievement distribution. The impact of instructional days on student achievement is monotonically increasing across the achievement distribution. These results suggest that the linear OLS, instrumental variables, and fixed effects estimates of the average impact of instructional days on academic achievement fail to recognize a nuanced, but policy-relevant feature of the distribution of instructional-day effects on achievement.

Specifically, the result that additional instructional time are more beneficial to higher achievers has implications for the growth and persistence of achievement gaps (Bitler et al., 2015) and is consistent with the theory that “skills beget skills.” Identifying the mechanisms underlying these results is outside the scope of the current study, but would be useful for future research to consider.

References

- Aucejo, E. M., Romano, T. F., 2013. Assessing the effect of school days and absences on test score performance. CFP Discussion Paper No. 1302, London School of Economics.
- Bellei, C., 2009. Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review* 28 (5), 629-640.
- Bitler, M., Domina, T., Penner, E., Hoynes, H., 2015. Distributional analysis in educational evaluation: A Case Study from the New York City Voucher Program. *Journal of Research on Educational Effectiveness*, 8 (3), 419-450.
- Cannon, J., Jacknowitz, A., Painter, G., 2006. Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis & Management* 25 (2), 299-321.
- Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., 2009. Are teacher absences worth worrying about in the U.S.? *Education Finance and Policy*, 4 (2), 115–149.
- DeCicca, P., 2007. Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review* 26 (1), 67-82.
- Eide, E., Showalter, M., 1998. The effect of school quality on student performance: A quantile regression approach. *Economic Letters* 58, 345-350.
- Fitzpatrick, M. D., Grissmer, D., Hastedt, S., 2011. What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review* 30 (2), 269-279.
- Gibbs, C., 2014. Experimental evidence on early intervention: The impact of full-day kindergarten. Unpublished manuscript, University of Virginia.
- Goodman, J., 2014. Flaking out: Student absences and snow days as disruptions of instructional time. National Bureau of Economic Research Working Paper No. w20221.
- Hansen, B., 2008. School year length and student performance: Quasi-experimental evidence. Unpublished manuscript, University of Oregon.
- Herrmann, M. A., Rockoff, J. E., 2012. Worker absence and productivity: Evidence from Teaching. *Journal of Labor Economics*, 30 (4), 749-782.
- Leuven, E., Lindahl, M., Oosterbeek, H., Webbink, D., 2010. Expanding schooling opportunities for four year olds. *Economics of Education Review* 29 (3), 319-328.

- Marcotte, D., Hemelt, S., 2008. Unscheduled school closings and student performance. *Education Finance and Policy* 3 (3), 316-338.
- Parinduri, R., 2014. Do children spend too much time in schools? Evidence from a longer school Year in Indonesia. *Economics of Education Review* 41, 89-104.
- Pischke, J., 2007. The impact of the length of school year on student performance and earnings: Evidence from the German short schooling years. *Economic Journal* 117 (523), 1216-42.
- Powell, D. 2015. Quantile regression with nonadditive fixed effects. Unpublished manuscript. Available at: http://works.bepress.com/david_powell/1
- Quinn, D., 2014. Black-white summer learning gaps: Interpreting the variability of estimates across representations. *Educational Evaluation and Policy Analysis* 37 (1), 50-69.
- Sims, D., 2008. Strategic responses to school accountability measures: It's all in the timing. *Economics of Education Review* 27 (1), 58-68.

Table 1
Descriptive Statistics

	Mean	SD
<i>Reading Achievement</i>		
Standardized fall K score	0.00	1.00
Standardized spring K score	1.18	1.22
Unadjusted standardized K school-year gain score	1.18	0.73
<i>Math Achievement</i>		
Standardized fall K score	0.00	1.00
Standardized spring K score	1.13	1.20
Unadjusted standardized K school-year gain score	1.13	0.69
<i>Calendar Days Between Important Dates</i>		
Days between K tests	187.3	21.9
<i>Student Characteristics</i>		
Black	17.4%	
Hispanic	13.9%	
Male	50.9%	
Poverty	17.4%	
Mom has high school diploma	29.8%	
Mom attended some college	31.9%	
Mom has bachelor's degree	14.7%	
Mom has more than a bachelor's degree	7.3%	
<i>Classroom Characteristics</i>		
Full Day Kindergarten	51.5%	
Classroom Size	20.7	5.0
Teacher's Age	37.3	16.2
Teacher has a graduate degree	32.1%	
Teacher's experience	8.4	7.9
Teacher is Black	6.2%	
Teacher is Hispanic	3.6%	
N Children	16,050	
N Schools	950	

Notes: Standard deviations (SD) are reported for non-binary variables. K = kindergarten. Achievement gains are not adjusted for differences in test dates.

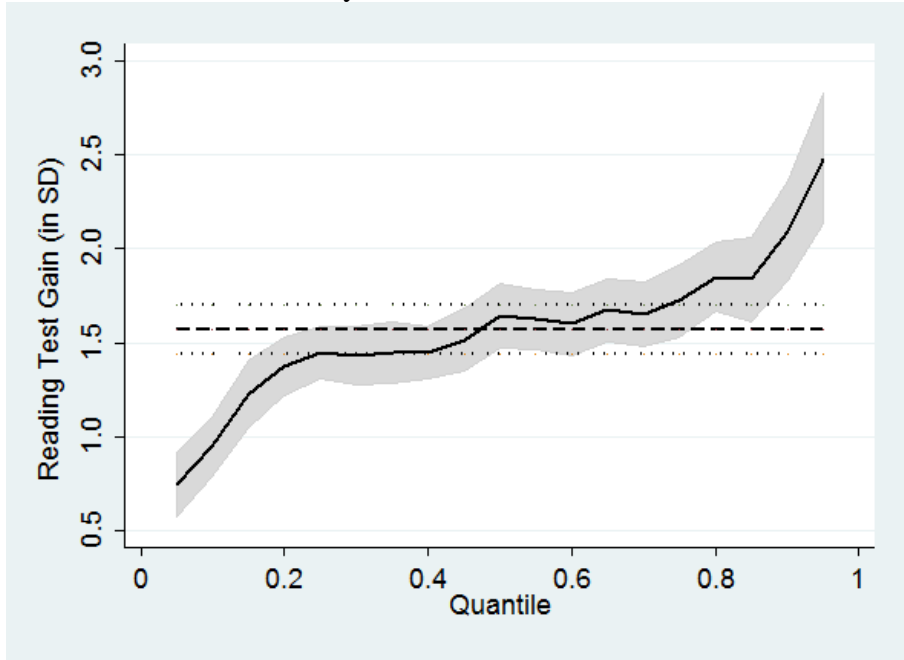
Table 2

OLS and quantile regression estimates of effect of 250-day period on test-score gains

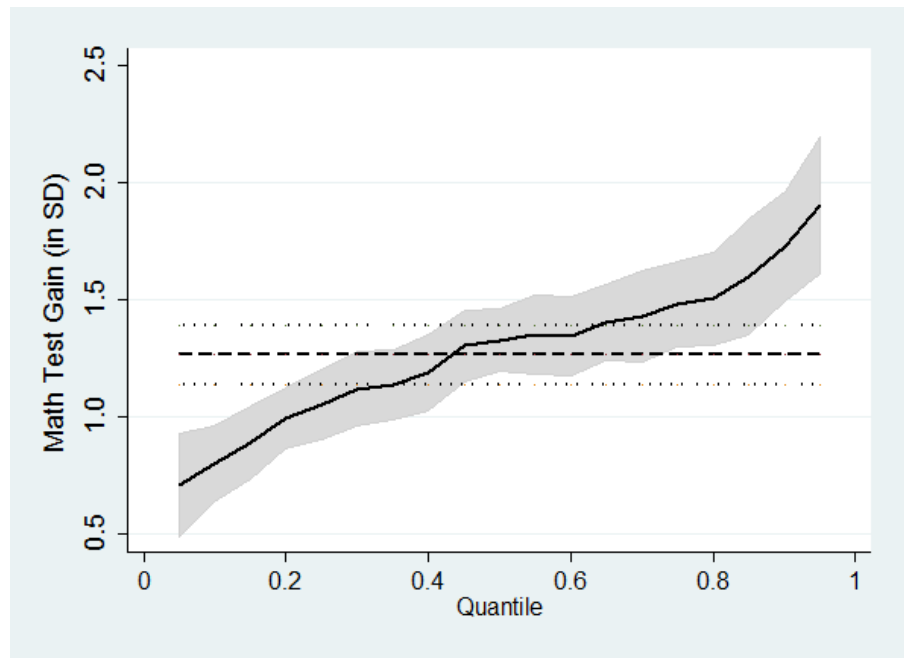
	Reading K Gain Score		Math K Gain Score	
	No Controls (1)	Controls (2)	No Controls (3)	Controls (4)
OLS	1.557*** (0.106)	1.617*** (0.103)	1.249*** (0.090)	1.321*** (0.086)
Quantile				
0.05	0.746*** (0.088)	0.809*** (0.109)	0.698*** (0.136)	0.643*** (0.140)
0.10	0.941*** (0.109)	1.016*** (0.100)	0.787*** (0.099)	0.835*** (0.094)
0.25	1.436*** (0.096)	1.442*** (0.104)	1.040*** (0.089)	1.114*** (0.085)
0.50	1.632*** (0.122)	1.651*** (0.109)	1.312*** (0.091)	1.392*** (0.101)
0.75	1.708*** (0.139)	1.841*** (0.133)	1.463*** (0.128)	1.522*** (0.135)
0.90	2.077*** (0.209)	2.113*** (0.190)	1.710*** (0.141)	1.799*** (0.141)
0.95	2.457*** (0.217)	2.535*** (0.212)	1.884*** (0.166)	2.020*** (0.177)

Notes: N = 16,050. K = kindergarten. IRT = item response theory. Standard errors in parentheses are robust to clustering at school level. Quantile regression standard errors were bootstrapped (500 replications). Other regressors include student characteristics (race, gender, poverty status, mother's education) and classroom characteristics (full-day K status, classroom size, teacher age, teacher experience, and teacher race). *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Estimates of Effect of 250 Day School Year on Student Achievement



A. Reading Achievement and No Controls



B. Math Achievement and No Controls

Notes: Dashed and solid lines represent linear and quantile regression coefficient estimates, respectively. Dotted lines and shaded areas represent corresponding 95% confidence intervals.

Online Appendix Materials

Online Appendix A

Weighted estimates of effects of 250-day units on academic achievement

	Reading K Gain Score		Math K Gain Score	
	No Controls	Controls	No Controls	Controls
OLS	1.468*** (0.110)	1.571*** (0.106)	1.248*** (0.097)	1.327*** (0.092)
Quantile				
0.05	0.734*** (0.102)	0.893*** (0.069)	0.822*** (0.133)	0.742*** (0.112)
0.10	0.799*** (0.090)	1.045*** (0.082)	0.847*** (0.100)	0.828*** (0.092)
0.25	1.361*** (0.094)	1.389*** (0.086)	1.060*** (0.085)	1.153*** (0.080)
0.50	1.560*** (0.101)	1.555*** (0.091)	1.306*** (0.090)	1.343*** (0.089)
0.75	1.687*** (0.108)	1.809*** (0.103)	1.398*** (0.114)	1.497*** (0.109)
0.90	1.970*** (0.160)	2.096*** (0.132)	1.746*** (0.151)	1.838*** (0.126)
0.95	2.203*** (0.210)	2.456*** (0.156)	2.025*** (0.174)	2.174*** (0.173)

Notes: N = 16,050. K = kindergarten. IRT = item response theory. These models are otherwise identical to the preferred baseline specification. Standard errors in parentheses are robust to clustering at school level. Quantile standard errors were bootstrapped (500 replications). Other regressors include student characteristics (race, gender, poverty status, mother's education) and classroom characteristics (full-day K status, classroom size, teacher age, teacher experience, and teacher race). *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix B
 Estimated effects of 250-day units on raw IRT scale scores

	Reading K Gain Score		Math K Gain Score	
	No Controls	Controls	No Controls	Controls
OLS	13.337*** (0.905)	13.855*** (0.882)	9.201*** (0.664)	9.728*** (0.630)
Quantile				
0.05	6.388*** (0.789)	6.935*** (0.926)	5.143*** (0.943)	4.737*** (0.971)
0.10	8.062*** (0.959)	8.702*** (0.858)	5.798*** (0.692)	6.148*** (0.721)
0.25	12.304*** (0.862)	12.354*** (0.831)	7.659*** (0.679)	8.206*** (0.608)
0.50	13.983*** (1.033)	14.147*** (0.897)	9.665*** (0.683)	10.250*** (0.719)
0.75	14.634*** (1.174)	15.775*** (1.149)	10.774*** (0.938)	11.212*** (0.994)
0.90	17.796*** (1.762)	18.108*** (1.691)	12.592*** (1.113)	13.250*** (1.000)
0.95	21.053*** (1.813)	21.720*** (1.865)	13.880*** (1.279)	14.876*** (1.335)

Notes: N = 16,050. K = kindergarten. IRT = item response theory. These models use raw IRT reading and math scale scores (means are 10.1 and 8.2, respectively), but are otherwise identical to the preferred baseline specification. Standard errors in parentheses are robust to clustering at school level. Quantile standard errors were bootstrapped (500 replications). Other regressors include student characteristics (race, gender, poverty status, mother's education) and classroom characteristics (full-day K status, classroom size, teacher age, teacher experience, and teacher race). *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix C

Lag-score estimates of 250-day units on academic achievement

	Reading K Gain Score		Math K Gain Score	
	No Controls	Controls	No Controls	Controls
OLS	1.562*** (0.107)	1.596*** (0.105)	1.247*** (0.092)	1.295*** (0.087)
Quantile				
0.05	0.744*** (0.098)	0.746*** (0.100)	0.601*** (0.125)	0.580*** (0.118)
0.10	0.924*** (0.113)	1.024*** (0.104)	0.694*** (0.096)	0.741*** (0.086)
0.25	1.394*** (0.105)	1.331*** (0.096)	1.031*** (0.087)	1.070*** (0.086)
0.50	1.634*** (0.124)	1.586*** (0.105)	1.316*** (0.097)	1.398*** (0.107)
0.75	1.746*** (0.144)	1.812*** (0.125)	1.485*** (0.143)	1.542*** (0.144)
0.90	2.089*** (0.176)	2.239*** (0.168)	1.726*** (0.158)	1.791*** (0.145)
0.95	2.468*** (0.214)	2.565*** (0.202)	1.885*** (0.195)	1.986*** (0.194)

Notes: N = 16,050. K = kindergarten. These lag-score models are otherwise identical to the preferred baseline gain-score specification. Standard errors in parentheses are robust to clustering at school level. Quantile standard errors were bootstrapped (500 replications). The dependent variable is the spring standardized test score. Other regressors include the fall standardized test score, student characteristics (race, gender, poverty status, mother's education) and classroom characteristics (full-day K status, classroom size, teacher age, teacher experience, and teacher race). *** p<0.01, ** p<0.05, * p<0.1.

Appendix D

Non-additive school FE quantile regression estimates of 250-day period on test-score gains

Quantile	Reading K Gain Score	Math K Gain Score
0.05	0.746*** (0.003)	0.698*** (0.001)
0.10	0.938*** (0.001)	0.781*** (0.002)
0.25	1.436*** (0.001)	1.040*** (0.003)
0.50	1.632*** (0.001)	1.310*** (0.002)
0.75	1.706*** (0.001)	1.463*** (0.001)
0.90	2.077*** (0.011)	1.709*** (0.001)
0.95	2.449*** (0.005)	1.885*** (0.007)

Notes: N = 16,050. FE = fixed effects. K = kindergarten. These non-additive school FE quantile regression models were estimated using the method proposed by Powell (2015).

*** p<0.01, ** p<0.05, * p<0.1.

Online Appendix E

Average Partial Effects (APEs) of 250-day units on academic achievement

	Reading K Gain Score			Math K Gain Score		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
OLS	1.570*** (0.107)	1.576*** (0.110)	1.538*** (0.103)	1.260*** (0.091)	1.270*** (0.089)	1.258*** (0.086)
Quantile						
0.05	0.746*** (0.088)	0.721*** (0.096)	0.732*** (0.105)	0.698*** (0.136)	0.701*** (0.137)	0.709*** (0.132)
0.10	0.941*** (0.109)	0.935*** (0.120)	0.911*** (0.102)	0.787*** (0.099)	0.790*** (0.105)	0.784*** (0.102)
0.25	1.436*** (0.096)	1.462*** (0.101)	1.400*** (0.106)	1.040*** (0.089)	1.050*** (0.089)	1.054*** (0.100)
0.50	1.632*** (0.122)	1.646*** (0.136)	1.575*** (0.120)	1.321*** (0.091)	1.311*** (0.100)	1.280*** (0.097)
0.75	1.708*** (0.139)	1.709*** (0.141)	1.656*** (0.138)	1.463*** (0.128)	1.453*** (0.145)	1.465*** (0.146)
0.90	2.077*** (0.209)	2.067*** (0.197)	2.053*** (0.194)	1.710*** (0.141)	1.723*** (0.139)	1.716*** (0.138)
0.95	2.457*** (0.217)	2.409*** (0.225)	2.367*** (0.214)	1.884*** (0.166)	1.893*** (0.182)	1.848*** (0.180)

Notes: N = 16,050. K = kindergarten. These models are otherwise identical to the preferred baseline specification. Standard errors in parentheses are robust to clustering at school level. Quantile standard errors were bootstrapped (500 replications).

*** p<0.01, ** p<0.05, * p<0.1