

## Spring or Fall Annual Tests? Implications for Value-added Models

Michael S. Hayes<sup>1</sup>, Rutgers University – Camden  
Seth Gershenson, American University and IZA

This manuscript is published in *Teachers College Record*. The appropriate citation is:

Hayes, Michael S., & Gershenson, Seth. 2018. Spring or fall annual tests? Implications for Value-added Models. Forthcoming, *Teachers College Record*, September, ID No. 22493.

### Abstract

School districts rely on standardized tests that are only administered once per academic year to produce value-added measures (VAMs) of teacher effectiveness. This is problematic because students' summer learning is incorrectly attributed to the teacher, potentially biasing estimates of teacher effectiveness. However, there is limited research on whether spring or fall tests yield more valid VAMs. We fill this gap in knowledge by comparing the accuracy of fall-to-fall and spring-to-spring "cross-year" VAMs relative to arguably more valid fall-to-spring "within-year" VAMs. We find that spring-to-spring "cross-year" VAMs, relative to fall-to-fall "cross-year" VAMs, are more valid, as they are more consistent with "within-year" VAMs. This suggests that spring assessments are preferred to fall assessments, at least when the objective is to obtain valid VAM-based estimates of school or teacher effectiveness.

Keywords: value-added models, teacher effectiveness, summer learning loss

---

<sup>1</sup> Corresponding author: Michael S. Hayes, Department of Public Policy & Administration, Rutgers University, 401 Cooper Street, Room 302, Camden, NJ 08102. Phone: 856-225-6561; Email: michael.hayes@rutgers.edu.

## 1. Introduction

Most school districts rely on standardized tests that are only administered once per academic year, usually in the spring, to evaluate teacher and school effectiveness. For example, the Nashville school district administers standardized tests each spring that measure students' achievement gains between the spring of grade  $g - 1$  and the spring of grade  $g$ . However, this approach is potentially problematic, as students' summer learning is incorrectly attributed to students' grade- $g$  teachers and schools. The resultant potential bias in value-added measures (VAMs) of teacher and school effectiveness when using such "cross-year" or "spring-to-spring" achievement gains is well documented (Downey et al., 2008; McEachin and Atteberry, 2017; Gershenson and Hayes, 2018).

One seemingly attractive solution is to administer tests at both the start and end of the school year and estimate VAMs on the associated within-year gains. However, this is not a panacea, as schools and teachers being evaluated on growth would have an incentive to artificially depress the fall scores. Moreover, adding a second round of tests would be costly, in terms of both time and money, not to mention the political costs of adding more tests at a time when many stakeholders are calling for fewer tests (Superville, 2015; Ujifusa, 2015). Accordingly, if high-stakes tests are to be administered only once per year, it is vital for school administrators to know whether implementing those assessments in spring or fall yields more credible VAMs.

The current study fills a gap in the VAM literature by directly addressing the question of whether fall or spring assessments yields more accurate VAMs. We do so by comparing the accuracy of fall-to-fall and spring-to-spring "cross-year" VAMs relative to the arguably more valid fall-to-spring "within-year" VAMs. Specifically, we use student-level data from the Early

Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K) to estimate both “cross-year” and “within-year” VAMs for first-grade classrooms and then compare the VAM-based rankings. We find that spring-to-spring “cross-year” VAMs, relative to fall-to-fall “cross-year” VAMs, are more reliable and consistent with “within-year” VAMs. This suggests that when only one assessment per year is feasible, spring assessments are preferred to fall assessments, at least when the objective is to obtain valid estimates of school or teacher effectiveness.

## **2. Data**

The current study uses student-level data from the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K) to estimate classroom valued-added scores. The ECLS-K, collected by the National Center for Education Statistics (NCES), is a longitudinal dataset comprised of a nationally-representative sample of the 2010-11 kindergarten cohort. The full sample contains approximately 18,000 children in over 900 kindergarten programs.<sup>1</sup> The survey oversampled certain subgroups of children, and the current study uses NCES-provided sampling weights to adjust for the survey’s nonrandom sampling frame. The main results remain qualitatively similar when the sampling weights are not applied, which is reassuring because it suggests that the main result is not driven by schools that have relatively higher number of disadvantaged students compared to the typical U.S. school (Solon, Haider, & Wooldridge, 2015).

The ECLS-K administered age-appropriate reading and mathematics tests to all surveyed children in the fall and spring of kindergarten and the spring of first grade.<sup>2</sup> The ECLS-K computed vertically scaled test scores using Item Response Theory (IRT). In the baseline

analysis, we use the unstandardized version of these IRT test scores. However, results are qualitatively similar if we standardize by subject, grade, and semester. We caution readers to avoid extrapolating the current study's results to other grade levels.

Our analysis also requires test score data in the fall of both first grade and second grade. However, only a random subsample of ECLS-K children were surveyed in the falls of first and second grades. Therefore, the analytic sample is restricted to this subsample, which includes approximately 3,700 children. The fall observations facilitate the following calculations:

- Test-score change between spring of kindergarten (K) and spring of first grade (1)
- Test-score change between fall of first grade (1) and spring of first grade (1)
- Test-score change between fall of first grade (1) and fall of second grade (2)

We make three additional sample restrictions. First, students who changed schools anytime in the years between kindergarten and second grade, or who experienced a midyear classroom change, are excluded. School changers are excluded to avoid conflating summer learning with disruptions to learning associated with changing schools (Schwartz et al., 2017). Second, we exclude students who repeated or skipped kindergarten, first grade, or second grade. Lastly, students are excluded if they were missing basic demographic data or classroom indicators. These additional restrictions yield a baseline analytic sample of approximately 1,800 first graders in over 750 classrooms. To check if the main results are not driven by imprecision associated with the relatively small number of students per classroom, we run the analysis with a subsample of 150 classrooms that have at least four surveyed students and find qualitatively similar results.

The ECLS-K is an ideal dataset for the current study for at least two reasons. First, the ECLS-K is the only nationally representative survey of U.S. students with test score data

spanning two summer vacations (i.e., summers after kindergarten and first grade) that links students to classrooms. This feature of the ECLS-K allows for the estimation of both “within-year” and “cross-year” VAMs for the same cohort’s first-grade school year. Second, the ECLS-K collects data from parents on students’ summer activities, which allows us to test if conditioning on summer activities reduces the bias inherent in “cross-year” VAMs. Online Appendix Table 1.A summarizes the student characteristics and summer activities of the analytic sample of ECLS-K first-graders.

Three features of the ECLS-K assessments require further explanation. First, due to the cohort nature of the ECLS-K, we only observe teachers in one school year; therefore, we can only identify “classroom” and not teacher effects. Second, the ECLS-K assessments were administered to different students on different days. A small number of ECLS-K administrators met individually with each student when administering an assessment, and this resulted in variation in test dates across schools, classrooms, and even students within the same classrooms (Fitzpatrick et al., 2011). Third, the assessments were not administered on the first or last days of the academic year. This is potentially problematic because some kindergarteners took the exam well in advance of the end of kindergarten and some first graders took the test well after the start of first grade. To account for this, we follow Quinn (2015) in adjusting for the timing of the test by extrapolating each math and reading test score to the first (or last) day of the school year.

For the summer between kindergarten (K) and first grade (1), and the summer between first grade and second grade (2), there are nine relevant dates ( $d$ ):  $d^{Fall,K}$ ,  $d^{Spring,K}$ ,  $d^{End,K}$ ,  $d^{Start,1}$ ,  $d^{Fall,1}$ ,  $d^{Spring,1}$ ,  $d^{End,1}$ ,  $d^{Start,2}$ , and  $d^{Fall,2}$ .<sup>3</sup> The extrapolations follow two steps. First, we calculate the daily learning rate during the relevant academic year for each child. Second, assuming that the same daily learning rate applies to the start and end of the year, we extrapolate

what test scores would have been at the beginning and end of the school year. For example, we use the following equations to calculate the extrapolated end of K test score:

$$Rate^K = \frac{(y^{Spring,K} - y^{Fall,K})}{(d^{Spring,K} - d^{Fall,K})} \quad (1)$$

$$y^{End,K} = y^{Spring,K} + [Rate^K * (d^{End,K} - d^{Spring,K})] \quad (2)$$

where  $y^j$  represents achievement at date  $j$  for  $j \in (d^{Fall,K}, d^{Spring,K}, d^{End,K})$ . Only the first two of these are observed. Equation (1) calculates the child-specific daily learning rate in K. Equation (2) uses the spring assessment score in kindergarten and the kindergarten daily learning rate to calculate the predicted assessment score at the end of kindergarten for each child. Table 1 reports both math and reading extrapolated test scores for all relevant dates. Interestingly, average math achievement appears to increase by almost 2 points between the end of kindergarten and the start of first grade. However, there is a slight decrease in average math achievement between the end of first grade and the start of second grade. Table 1 also shows that the average child experiences summer reading loss over both summers.

### 3. Methods

We utilize the ECLS-K data discussed in Section 2.1 to make three sets of comparisons. First, we compare VAM-based rankings of classroom effectiveness generated by fall-to-spring (within-year) achievement gains to the arguably less-valid rankings generated by spring-to-spring (cross-year) gains; the corresponding VAM specifications are given by Equations 3a and 3b, respectively:

$$y_{ic}^{End,1} = \gamma + \alpha y_{ic}^{Start,1} + \beta x_i^1 + \theta_c^1 + u_{ic} \quad (3a)$$

and

$$y_{ic}^{End,1} = \gamma + \alpha y_{ic}^{End,K} + \beta x_i^1 + \theta_c^1 + u_{ic}. \quad (3b)$$

Second, we compare VAM-based rankings of classroom effectiveness generated by Equation (3a) to the arguably less-valid rankings generated by fall-to-fall (cross-year) gains shown in Equation (3c) below:

$$y_{ic}^{Start2} = \gamma + \alpha y_{ic}^{Start,1} + \beta x_i^1 + \theta_c^1 + u_{ic}. \quad (3c)$$

Lastly, we compare VAM-based rankings from the two cross-year equations (3b and 3c).

In Equation 3, student and classrooms are indexed by  $i$  and  $c$ , respectively;  $K$ , 1, and 2 indicate kindergarten, first grade, and second grade, respectively;  $y$  is academic achievement (i.e., extrapolated math and reading scores); vector  $x$  contains some combination of the student characteristics and summer activities described in Appendix Table 1.A;  $\theta$  are the classroom fixed effects (FE) upon which rankings of classroom effectiveness will be based; and  $u$  is a mean-zero error term that captures the unobserved predictors of achievement. All equations are estimated by ordinary least squares (OLS). The baseline model contains only a limited set of the student demographic variables including indicators for race, gender, poverty, English language learner (ELL) status, individualized education plan (IEP) status, kindergartener redshirt status, attending private school, attending an urban school, and attending a rural school (Gershenson and Hayes, 2017). As a robustness check, we add additional controls for student demographic characteristics and summer activities to all models. We compare the rankings generated by cross- and within-year VAMs in two ways similar to previous researchers (Gershenson and Hayes, 2018; Guarino, Reckase, & Wooldridge, 2015; Koedel and Betts, 2007; Loeb and Candelaria, 2012; McCaffrey et al., 2009; McEachin and Atteberry, 2017). First, we estimate Spearman Rank Correlations, which are simple summary statistics that measure the similarity between two rankings. Second, we construct transition matrixes that document switching across specifications, which provide a more nuanced understanding of how the rankings change and of the implications for policies that

penalize (reward) teachers in the bottom (top) of the effectiveness distribution.

#### **4. Results**

Table 2 reports Spearman rank correlations of the comparisons between estimated first-grade classroom effects generated by Equations 3a, 3b, and 3c for the baseline specification as well as several alternative specifications. The Spearman rank correlations suggest that estimated classroom effects from VAMs using spring-to-spring achievement gains for both subjects are more robust to test timing than similar VAMs using fall-to-fall achievement gains. In fact, for both math and reading achievement, the Spearman rank correlation coefficients are more than 10 percentage points higher when using the classroom effects generated by spring-to-spring achievement gains relative to the fall-to-fall achievement gains. Not surprisingly, shown in Column 3, the Spearman rank correlation coefficients are the smallest when comparing the rankings of classroom effects generated from VAMs using fall-to-fall achievement gains to similar classroom effects from spring-to-spring achievement gains.

The results reported in Table 2 are robust to changes to the baseline specification. For example, removing or adding control variables on student characteristics and summer activities do not appreciably change the Spearman correlation coefficients. This finding is not surprising as previous research suggests that only 10% of the variation of summer learning can be explained by student and household characteristics (Downey et al., 2004), and also the data on summer activities in the ECLS-K do not contain detailed information on the quality of summer activities and parent involvement over the summer. The result is also robust to excluding the NCES-provided sampling weights mentioned in Section 2. Similarly, the main result is robust to including only classrooms with at least four surveyed students, which suggest that the findings

are not driven by imprecision associated with the relatively small number of students per classroom.

Table 3 presents transition matrixes for math and reading achievement based on the baseline value-added model that conditions on elements of  $\mathbf{x}$  typically observed in administrative data. Transition matrixes report the movement of classrooms across quintiles of the classroom-effectiveness distribution, which provides a more nuanced understanding of the stability of the rankings reported in Table 2. The diagonal elements of the transition matrixes reported in Table 3 represent classrooms that were in the same quintile of the effectiveness rankings generated by fall-to-fall and fall-to-spring VAMs. As expected given the results in Table 2, the figures along the diagonals are significantly lower than 100%, reinforcing the general finding that 1<sup>st</sup> grade classroom-effectiveness rankings are sensitive to the timing of the assessments used in the VAM. Indeed, only about half of classrooms ranked in the lowest or highest quintiles of math effectiveness remained in the same quintile in both the within-year and cross-year rankings.

Table 4 replicates the transition matrix analysis in Table 3 for specifications (Equation 3a) and (Equation 3b), comparing the rankings of first-grade classrooms generated by fall to spring VAMs to those generated by spring-to-spring VAMs. Table 4 shows that the spring-to-spring VAMs are more stable than the fall-to-fall VAMs, and large swings cross multiple quintiles are exceedingly rare in both subjects. Overall, the results from Table 4 are consistent with the main finding, from Table 2, that estimated classroom effects from VAMs using spring-to-spring achievement gains for both subjects are more robust to test timing than similar VAMs using fall-to-fall achievement gains.

## **5. Discussion**

The current study addressed a common problem facing the majority of U.S. school districts: the difficulties of estimating teacher effectiveness with standardized tests administered only once per year. The problem is that students' summer learning gains and losses are incorrectly attributed to schools and teachers when cross-year VAMs are used to evaluate teacher effectiveness. Indeed, previous research has documented this potential bias in cross-year VAM estimates (Downey et al., 2008; McEachin and Atteberry, 2017; Gershenson and Hayes, 2018). Given the political and financial challenges to administering standardized exams twice per year, school administrators need to be aware of whether implementing those assessments in spring or fall yields more credible VAMs. Ours is the first study to directly address this question.

The current study provides evidence on the validity of value-added estimates of classroom effects generated by fall-to-fall and spring-to-spring "cross-year" VAMs relative to arguably more valid fall-to-spring "within-year" VAMs. We consistently find that estimated classroom effects from VAMs using spring-to-spring achievement gains for both subjects are more robust than similar VAMs using fall-to-fall achievement gains. Specifically, for both math and reading achievement, the Spearman rank correlation coefficients are more than 10 percentage points higher when using the classroom effects generated spring-to-spring achievement gains relative to the fall-to-fall achievement gains. Transition matrixes reported in the current study provide a similar finding. The policy implication from this finding is that when only one assessment per year is feasible, spring assessments are preferred to fall assessments, at least when the objective is to obtain valid estimates of school or teacher effectiveness. Moving forward, one area for future research is determining the optimal timing of the spring assessment. Assuming the validity of the spring-to-spring "cross-year" VAMs increases monotonically as the school year progresses, we would predict that a spring-to-spring "cross-year" VAM administered

in March would be more valid than a similar spring-to-spring “cross-year” VAM administrated in February, but less valid than one administrated in April. Therefore, if one is willing to make this sort of monotonicity assumption, we can say that tests administered later in the school year are better than tests administered earlier in the year. Unfortunately, our current data does not allow us to formally test this hypothesis.

## Notes

1. Sample sizes are rounded to the nearest 50, as per NCES rules for restricted-use ECLS-K data.
2. See Fitzpatrick et al. (2011) and Quinn (2014) for more discussion of the tests.
3. Unfortunately, the ECLS-K does not report the exact day of the assessment. Instead, an indicator for the week the assessment was administered is provided. We impute test dates by converting the week indicators to the midpoint of the week (e.g., if week 1 covers the 1<sup>st</sup> through the 7<sup>th</sup>, we impute a test date of the 4<sup>th</sup>). This should not create any systematic bias due to the conditional randomness of the test dates (Fitzpatrick et al., 2011).

## References

- Gershenson, S., Hayes, M.S. (2017). The summer learning of exceptional students. *American Journal of Education*, 123(3), 447-473.
- Gershenson, S., Hayes, M.S. (2018). The implications of summer learning loss for value-added estimates of teacher effectiveness. *Educational Policy*, 32(1), 55-85.
- Downey, D.B., von Hippel, P.T., & Broh, B. (2004). Are schools the great equalizer? Cognitive Inequality during the summer months and the school year. *American Sociological Review*, 69(5), 613-635.
- Guarino, C.M., Reckase, M.D., & Wooldridge, J.M. (2015). Can value-added measures of teacher performance be trusted? *Education Finance & Policy*, 10(1), 117-156.
- 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.
- Koedel, C., & Betts, J.R. (2011). Does student sorting invalidate value-added models of teacher Effectiveness? An extended analysis of the Rothstein critique. *Education Finance & Policy*, 6(1), 18-42.
- Loeb, S., & Candelaria, C. (2012). How stable are value-added estimates across years, subjects, and student groups? The Carnegie Knowledge Network, Stanford, CA. Retrieved from [http://www.carnegieknowledge.org/wp-content/uploads/2012/10/CKN\\_2012-10\\_Loeb.pdf](http://www.carnegieknowledge.org/wp-content/uploads/2012/10/CKN_2012-10_Loeb.pdf)
- McCaffrey, D.F., Sass, T.R., Lockwood, J.R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance & Policy*, 4(4), 572-606.
- McEachin, A., & Atteberry, A. (2017). The impact of summer learning loss on measures of school performance. *Education Finance & Policy*, 12(4), 468-491.
- Quinn, D.M. (2015). Black-white summer learning gaps: Interpreting the variability of estimates Across representations. *Educational Evaluation and Policy Analysis*, 37(1), 50-69.
- Schwartz, A.E., Stiefel, L., & Cordes, S.A. (2017). Moving matters: The causal effect of moving schools on student performance. *Education Finance & Policy*, 12(4), 419-446.
- Superville, D.R. (2015). Students take too many redundant tests, study finds. Education Week.
- Ujifusa, A. (2015). Standardized testing costs states \$1.7 billion a year, study says. Education Week.
- Solon, G., Haider, S.J., & Wooldridge, J.M. (2015). What are we weighting for? *Journal of*

*Human Resources*, 50(2), 301-316.

**Table 1.** Descriptive Statistics of Extrapolated Test Scores

	Mean	SD
Math Extrapolated Test Scores		
End of Kindergarten	50.3	11.3
Start of 1 <sup>st</sup> Grade	52.0	14.5
End of 1 <sup>st</sup> Grade	72.4	14.0
Start of 2 <sup>nd</sup> Grade	71.9	13.6
Reading Extrapolated Test Scores		
End of Kindergarten	66.9	13.2
Start of 1 <sup>st</sup> Grade	66.5	15.9
End of 1 <sup>st</sup> Grade	91.1	13.9
Start of 2 <sup>nd</sup> Grade	89.8	13.1
N Students	1,800	
N Classrooms	750	
N Schools	250	

Notes: All ECLS-K estimates are weighted to account for the unequal probabilities of sample selection by NCES-provided sampling weights. ECLS-K sample sizes are rounded to the nearest 50, in accordance with NCES regulations for restricted-use ECLS-K data.

**Table 2: Spearman Correlation Coefficients for Ranking Comparisons**

	First-Grade Classrooms		
	Fall-Spring vs. Fall-Fall	Fall-Spring vs. Spring-Spring	Fall-Fall vs. Spring-Spring
<b>Math Achievement</b>			
No controls	0.66	0.80	0.50
Baseline	0.64	0.80	0.46
Rich control set	0.60	0.81	0.46
Baseline, unweighted	0.63	0.80	0.47
Baseline, restricted sample	0.61	0.79	0.45
<b>Reading Achievement</b>			
No controls	0.74	0.88	0.67
Baseline	0.73	0.88	0.67
Rich control set	0.71	0.89	0.64
Baseline, unweighted	0.73	0.88	0.66
Baseline, restricted sample	0.70	0.88	0.62
Students		1,800	
Classrooms		750	

Notes: The baseline model contains only a limited set of the student demographic variables including indicators for race, gender, poverty, English language learner (ELL) status, individualized education plan (IEP) status, kindergartener redshirt, attending private school, attending an urban school, and attending a rural school. The rich control set specification contains on variables summarized in Appendix Table A.1. All spearman correlation coefficients reported in Table 2 are strongly statistically significant with a p-value less than 0.0005. The restricted sample includes classrooms with at least four surveyed students. The restricted sample includes 700 students in 150 classrooms.

**Table 3.** Quintile Transitions of First-Grade Classroom Effects Generated by Fall-Spring and Fall-Fall Gains

Fall-spring	Fall-fall, baseline model				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Math Achievement					
Quintile 1	57.6	21.9	14.6	3.3	2.6
Quintile 2	27.8	30.4	19.9	17.9	3.9
Quintile 3	8.6	23.2	29.1	23.2	15.9
Quintile 4	4.6	14.6	25.8	31.8	23.2
Quintile 5	1.3	10.0	10.7	24.0	54.0
Reading Achievement					
Quintile 1	70.2	18.5	6.6	3.3	1.3
Quintile 2	20.5	43.0	22.5	10.6	3.3
Quintile 3	8.6	18.5	33.1	26.5	13.2
Quintile 4	0.0	12.6	25.2	38.4	23.8
Quintile 5	0.7	7.3	12.7	21.3	58.0

Notes: The statistics reported in this table compare rankings of classroom effects generated by Equations 3a and 3c of the main text. The sample contains 1,800 students and 750 classrooms.

**Table 4.** Quintile Transitions of First-Grade Classroom Effects Generated by Fall-Spring and Spring-Spring Gains

Fall-spring	Spring-spring, baseline model				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Math Achievement					
Quintile 1	68.2	24.5	6.0	0.6	0.6
Quintile 2	24.5	30.5	29.1	11.9	4.0
Quintile 3	6.0	32.5	30.5	21.9	9.3
Quintile 4	1.3	9.9	26.5	40.4	21.9
Quintile 5	0.0	2.7	8.0	25.3	64.0
Reading Achievement					
Quintile 1	80.1	17.2	2.0	0.0	0.7
Quintile 2	18.5	49.7	21.9	8.6	1.3
Quintile 3	1.3	27.8	41.7	25.2	4.0
Quintile 4	0.0	5.3	28.5	45.7	20.5
Quintile 5	0.0	0.0	6.0	20.7	73.3

Notes: The statistics reported in this table compare rankings of classroom effects generated by Equations 3a and 3b of the main text. The sample contains 1,800 students and 750 classrooms.

**Appendix Table A.1.** ECLS-K Sample’s Demographic Composition and Summer Activities

	Mean	SD
<i>Student Characteristics</i>		
White	57.6%	
Black	10.5%	
Hispanic	22.0%	
Other race/ethnicity	9.9%	
Female	49.6%	
Poverty	21.7%	
Does not speak English at Home	13.6%	
Has Individualized Education Plan (IEP)	7.0%	
Kindergarten Redshirt	9.0%	
Attends Private School	7.7%	
Attends Urban School	28.1%	
Attends Suburban School	42.6%	
Attends Rural School	29.4%	
Mom No H.S. Degree	10.2%	
Mom H.S. Degree	23.8%	
Mom Some College	33.4%	
Mom Bachelor’s Degree or more	32.6%	
<i>Summer Activities</i>		
Organized summer activities	97.2%	
Attended summer school	7.7%	
# of trips to library/bookstore	6.1	7.7
Child tutored over summer	5.4%	
Child received special services	1.5%	
Child attended day/overnight camp	28.8%	
Child never practice math	7.0%	
Child sometimes practices math	85.0%	
Child practices math everyday	8.0%	
Mother never reads to child	8.1%	
Mother sometimes reads to child	54.0%	
Mother reads to child everyday	37.8%	
N Children	1,800	
N Classrooms	750	
N Schools	250	

Notes: Means are weighted by NCES provided sampling weights to account for unequal probabilities of sample selection. Standard deviations are reported for non-binary variables.