

# Linking Teacher Quality, Student Attendance, and Student Achievement

[Forthcoming in *Education Finance and Policy*]

Seth Gershenson  
Assistant Professor of Public Policy  
School of Public Affairs  
American University  
4400 Massachusetts Avenue, NW  
Washington, DC 20016-8070

Email: [gershens@american.edu](mailto:gershens@american.edu)  
Phone: (202) 885-2687  
Fax: (202) 885-2347

## ABSTRACT

Research on the effectiveness of educational inputs, particularly research on teacher effectiveness, typically overlooks teachers' potential impact on behavioral outcomes such as student attendance. Using longitudinal data on teachers and students in North Carolina I estimate teacher effects on primary school student absences in a value-added framework. The analysis yields two main findings: First, teachers have arguably causal, statistically significant effects on student absences that persist over time. Second, teachers who improve test scores do not necessarily improve student attendance, suggesting that effective teaching is multidimensional and teachers who are effective in one domain are not necessarily effective in others.

JEL Codes: I21, J44, J45

Keywords: student absences, teacher effectiveness, value added models, non-cognitive skills

Acknowledgements: The author is grateful for financial support from the Spencer Foundation and the American Educational Research Association (AERA). AERA receives funds for its "AERA Grants Program" from the National Science Foundation under NSF Grant #DRL-0941014. Opinions reflect those of the author and not necessarily those of the funding agencies. The author thanks the North Carolina Education Research Data Center (NCERDC) for providing access to the restricted-use North Carolina data. The author thanks Nora Gordon, Cassie Guarino, Mike Hansen, two anonymous referees, seminar participants at American University, Johns Hopkins University, Oregon State University, The College Board, and The University of Oregon, and conference participants at the 2014 meetings of the Association for Education Finance & Policy and Society for Research on Educational Effectiveness for providing helpful comments. Andrew Brannegan and Michael S. Hayes provided excellent research assistance. Any remaining errors are my own.

## 1. Introduction

Research on the technology of skill formation routinely finds evidence of a direct causal relationship between character skills and long-run socioeconomic outcomes (Cunha, Heckman, and Schennach 2010; Heckman, Stixrud, and Urzua 2006).<sup>1</sup> For example, character skills such as conscientiousness, motivation, and self-discipline predict important socioeconomic outcomes such as educational attainment, employment, earnings, marriage, and crime (e.g., Almlund et al. 2011; Borghans et al. 2008; Heckman and Kautz 2013; Jackson 2013; Jacob 2002; Lundberg 2012, 2013). Attendance is an objectively measurable behavior that is correlated with at least three of the “Big Five” character skills identified by psychologists: Attendance is positively associated with conscientiousness (Duckworth et al. 2007) and negatively associated with neuroticism and low levels of agreeableness (Lounsbury et al. 2004).<sup>2</sup> Conscientiousness is a character skill that is valued in the labor market (Heckman and Kautz 2013) and regular attendance is highly valued by employers (Lerman 2013; Morrison et al. 2011; Pritchard 2013). Similarly, regular school attendance is positively associated with academic achievement (Aucejo and Romano 2013; Gershenson, Jackowitz, and Brannegan 2014; Gottfried 2009) and negatively associated with grade retention (Nield and Balfanz 2006), drug use (Hallfors et al. 2002), and dropping out of school (Rumberger and Thomas 2000).

As a result, identifying the educational inputs and interventions that improve students’ attendance is likely of interest to both educators and policy makers. However, most interventions are designed to increase cognitive skills, as measured by standardized tests, and evaluated

---

<sup>1</sup> Character skills encompass a variety of skills and behaviors that have previously been referred to as non-cognitive skills, non-cognitive ability, soft skills, character traits, personality traits, and socio-behavioral skills, among other names (Heckman and Kautz 2013).

<sup>2</sup> The Big Five character skills are Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN) (Heckman and Kautz 2013, pp. 10-12).

accordingly (Heckman 2000). This is despite the facts that character skills are more malleable than cognitive skills (Cunha and Heckman 2008; Heckman 2000) and that such a focus may cause teachers and schools to divert resources away from non-tested skills (Baker et al. 2010; Harris 2011). A similar critique applies to the large literature on teacher effectiveness, despite widespread agreement that teachers are the most important school-provided educational input (e.g., Clotfelter, Ladd, and Vigdor 2007; Goldhaber 2007; Harris 2011; Rivkin, Hanushek, and Kain 2005) and the fact that teachers likely affect students' development in numerous areas outside the reading and math skills measured by standardized tests (Ladd and Sorensen, 2014).

The current study contributes to this gap in the literature by estimating teacher effects on primary school students' absences in a value-added (VA) framework. This work complements research by Jackson (2013) on 9<sup>th</sup>-grade teachers' effects on an index of non-cognitive skills, as at least some of the mechanisms through which teachers affect primary school attendance likely differ from the ways that teachers affect secondary school attendance.<sup>3</sup> Moreover, identifying the educational inputs that improve the attendance of younger students is particularly important given that character skills are shaped by children's early environments (Heckman et al. 2006), problems of chronic absence and school disengagement manifest as early as first grade (Alexander, Entwisle, and Kabbani 2001), and socioeconomic gaps in character skills exist prior to kindergarten and grow over time (Duncan and Magnuson 2011). Improving the character skills and attendance habits of disadvantaged children will likely foster socioeconomic mobility and social inclusion and increase the returns to subsequent educational attainment (Heckman and Kautz 2013). The mechanisms through which teachers might affect primary school students' attendance are discussed in section 2.

---

<sup>3</sup> Student absences, suspensions, grade promotion, and grade point average (GPA) comprise the non-cognitive index used in Jackson (2013).

In addition to identifying what effect, if any, primary-school teachers have on student absences, the current study also contributes to the literature on the validity of VA estimates of teacher effectiveness more generally by addressing one of the central questions regarding VA articulated by Chetty, Friedman, and Rockoff (2013): Do high-VA teachers improve student outcomes other than test scores? Estimating teachers' effects on an objective outcome such as student absences addresses the common criticism that VA measures of teacher effectiveness focus too narrowly on students' performance on standardized tests. The focus on standardized tests is potentially problematic for several reasons: it may cause teachers and schools to divert resources away from non-tested topics and skills (Baker et al. 2010; Harris 2011), it disregards Fenstermacher and Richardson's (2005) broad definition of quality teaching, and it potentially biases estimates of teacher quality by ignoring teachers' effects on students' character skills and related behaviors (e.g., attendance, study habits, etc.) (Heckman 2000). Accordingly, I assess the importance of objectively evaluating teachers along multiple dimensions by comparing rankings of teacher effectiveness based on teachers' effects on test scores to similar rankings based on teachers' effects on student attendance. Significant differences between the two rankings would suggest that teacher evaluations based solely on teachers' abilities to improve student test scores miss an important dimension of teacher quality, systematically misclassify effective teachers as ineffective, and vice versa. Estimates of teachers' effects on student absences also provide objective measures of effectiveness for teachers who do not teach in tested grades or subjects.

Specifically, I address two research aims. First, I estimate teachers' effects on student attendance by estimating VA models that consider student attendance as an output of the education production function. Second, I estimate corresponding teacher effects on academic achievement (i.e., test scores) and compare the resulting rankings of teacher effectiveness to

rankings based on teachers' effects on student attendance. Both sets of VA models are estimated using rich longitudinal administrative data on both teachers and students from North Carolina. The main results generally suggest that teachers significantly affect student absences and that this relationship is arguably causal. Interestingly, teacher effectiveness is not stable across domains, as rank correlations between teachers' effects on test scores and teachers' effects on student absences are generally close to zero. Additional analyses show that these results are not specific to North Carolina, teachers' effects on student absences persist over time, and teachers' effectiveness in reducing absences is positively correlated over time and with teaching experience.

The paper proceeds as follows: Section 2 describes the mechanisms through which teachers might affect student attendance and briefly reviews the relevant literature on teacher effectiveness. Section 3 describes the data and Section 4 describes the identification strategy. Section 5 presents the main results and Section 6 presents further analyses of the intertemporal stability, persistence, and relationship with teaching experience of teachers' effects on student absences. Section 7 examines the cross-domain stability of teacher effectiveness by comparing rankings based on teachers' effects on student absences to rankings based on their effects on student achievement. Section 8 concludes.

## **2. Conceptual Background and Previous Literature**

Chetty et al. (2011) found small transitory effects of kindergarten classrooms on cognitive development (i.e., test scores) but significant effects on long-run outcomes such as earnings. One interpretation of these seemingly contradictory results is that teachers affect long-run outcomes by building students' non-cognitive skills (Jackson 2013). Indeed, Jackson (2013)

develops a formal latent factor model in which both student and teacher ability are two dimensional (i.e., cognitive and non-cognitive), and shows that teachers who affect students' non-cognitive development but not cognitive development can substantively affect students' long run outcomes. It is generally believed that instruction can improve character skills and there is a long history of using observed behaviors as proxies for character skills (Almlund et al. 2011; Heckman and Kautz 2013). Attendance is one such proxy, which is both objective and easily observable, that previous researchers have utilized (e.g., Jacob 2002; Jackson 2013).

Teachers potentially increase student attendance through some combination of fostering a passion for learning, increasing student engagement, creating a strong sense of community in the classroom, and stressing the importance of regular attendance (Baker et al. 2010; Kelly 2012; Ladd and Sorensen 2014; Monk and Ibrahim 1984). Of course, some of these mechanisms might be more relevant to older students whose attendance is arguably less influenced by their parents. Another way that elementary school teachers might affect young children's attendance is by influencing parents' and other household adults' attitudes towards children's school attendance and punctuality, as parental involvement is thought to be malleable.<sup>4</sup> Teachers might do so early in the school year at "back to school" nights or during parent-teacher conferences throughout the year. Moreover, anecdotal evidence from private conversations with primary school teachers suggests that some teachers initiate contact with students' parents in response to frequent absences. Some schools even have formal policies regarding parental outreach in response to student absences. For example, section 4400.4 of North Carolina's Newlin Elementary School's

---

<sup>4</sup> For example, 6% of respondents in a 2004 Gallup Poll listed "increasing parental involvement" as the "best way to improve K-12 education in the United States (U.S.)" (Gallup 2004).

2013-14 Parent/Student Handbook states that the school will initiate a student-parent conference after a student accumulates six unexcused absences.<sup>5</sup>

Teachers likely vary in their influence on non-cognitive behaviors such as attendance for at least three reasons. First, some teachers may simply be better than others at influencing students' character skills and/or parental involvement. Second, teachers' attitudes towards the importance of teaching character skills relative to academic skills may vary (Dombkowski 2001), resulting in differences across classrooms in time and effort allocations. Third, teachers may allocate effort based on their perceived ability to influence students' character skills, regardless of the importance they attach to influencing such skills (Jennings and DiPrete 2010).

To date, however, only four studies have empirically investigated the impact of teachers on students' character skills.<sup>6</sup> First, Dobbie (2011) found that some of the criteria used to determine admission into the Teach for America (TFA) program are associated with improved classroom behavior but little evidence of an effect of the TFA criteria on student absences. Second, Jennings and DiPrete (2010) found that kindergarten and first-grade teachers in the ECLS-K have sizable effects on a "social-behavioral index" that measures children's approaches to learning, self control, and interpersonal skills. Interestingly, the authors found that the teachers who had the largest effects on children's behavior did not always have large effects on children's test scores, suggesting that by focusing only on teachers' effects on test scores, effective teachers may be misclassified as ineffective, and vice versa. Third, Ladd and Sorensen (2014)

---

<sup>5</sup><http://www.abss.k12.nc.us/cms/lib02/NC01001905/Centricity/Domain/166/Newlin%20ElementaryHandbook%2013-14.pdf>

<sup>6</sup> The Talent Development high school program is a notable intervention designed to improve student attendance. Initially launched in five Philadelphia public high schools, the program increased student attendance by three to seven percent in the first three treated cohorts (Kempel et al. 2005). The program provided students with individualized support that, among other things, prioritized high attendance.

investigated the relationship between North Carolina middle school teachers' experience and student absences, time spent reading for pleasure, time spent on homework, and disruptive behavior in the classroom. The authors find significant effects of teacher experience on student absences. Finally, using administrative data from North Carolina, Jackson (2013) found that 9th-grade teachers have significant effects on students' non-cognitive skills, as measured by an index of student absences, suspensions, grade promotion, and grade point averages. Like Jennings and DiPrete, Jackson finds that many of the teachers who most effectively develop students' non-cognitive skills have only average effects on test scores, suggesting that focusing on test scores alone will fail to identify some effective teachers. These findings are consistent with the robust result in the VA literature that rankings of teacher effectiveness are not perfectly correlated across academic subjects (e.g., Koedel and Betts 2007; Goldhaber, Cowan, and Walch 2013; Lockwood et al. 2007; Loeb and Candelaria 2012; Loeb et al. 2012), though cross-subject rank correlations tend to be significantly more stable than the cross-domain rank correlations identified in the current study.

The general lack of attention paid to teachers' impacts on students' character skills is therefore surprising, as identifying effective teachers is hugely important and a growing consensus agrees that providing high-quality teachers to all students must play a prominent role in closing achievement gaps between students of different demographic and socioeconomic backgrounds (Harris 2011; Rivkin, Hanushek, and Kain 2005). VA models that attempt to identify individual teachers' contributions to gains in student achievement are gaining popularity and acceptance as useful measures of teacher effectiveness, though such measures remain controversial (Baker et al. 2010; Chetty et al. 2013; Harris 2011). Specifically, critics of VA measures of teacher effectiveness question whether policies that incentivize schools and teachers

to increase test scores displace beneficial classroom activities that develop character skills and learning in non-tested academic subjects. In addition to identifying primary-school teachers' effects on student attendance, the current study also contributes to the general VA literature by speaking to the practical significance of this criticism.

I assess the practical importance of the criticism that VA models focus too narrowly on students' performance on standardized tests by comparing the stability of teacher rankings of their effects on achievement gains to corresponding rankings based on teachers' effects on absences. If some teachers who excel at increasing test scores are less able to promote attendance, and vice versa, policies that evaluate teachers on only one dimension will necessarily misclassify a nontrivial subset of teachers. This idea is formalized in figure 1, which assumes that teacher quality is two dimensional. Teacher A is unambiguously the most effective teacher in figure 1, as teacher A exerts the largest impact on students' attendance and academic achievement. Note that if all two-dimensional measures of teachers' effectiveness were to fall on the dashed 45 degree line the dimension along which teachers are evaluated would not matter. Previous research suggests that this is not the case (Jackson 2013; Jennings and DiPrete 2010).

Now consider the effectiveness of teachers B, C, and D in figure 1. In the two-dimensional setting, teacher D is unambiguously the least effective and has the smallest impacts on both attendance and achievement. Meanwhile, teacher B excels at improving students' attendance while teacher C excels at improving students' academic achievement. However, an accountability system that evaluated teachers solely based on their ability to improve student test scores would mistakenly conclude that teacher C is more effective than teachers B and D, who appear equally effective. By estimating teachers' effects on both student absences and academic achievement, the current study identifies the ability of an important educational input (teachers)

to affect an important non-cognitive behavior (attendance). More generally, the current study provides evidence on the extent to which teachers excel along multiple dimensions and the general importance of evaluating teachers along multiple objective dimensions.

### **3. Data**

I estimate teachers' effects on student absences using longitudinal administrative data on the population of third through fifth graders who attended North Carolina's public schools between the 2005-06 and 2009-10 school years. These student-level data are maintained and provided by the North Carolina Education Research Data Center (NCERDC).<sup>7</sup> The NCERDC data contain administrative records on students' race, gender, poverty status, limited English proficiency (LEP) status, whether the student had administratively classified math or reading learning disabilities, total absences, student-teacher links, and end-of-grade math and reading test scores.<sup>8</sup> North Carolina's end-of-grade tests are state mandated, criterion referenced, vertically aligned, and are given to all students in the spring of third, fourth, and fifth grades. Third-grade and 2006 data are used as lags in value-added models and thus the analytic sample is comprised of fourth and fifth graders between 2007 and 2010. Students who either experienced a mid-year classroom change, repeated third, fourth, or fifth grade, or are missing achievement, absence, or demographic data are excluded from the analysis. These exclusions result in an analytic sample of 446,244 student-year observations, 27,943 unique classrooms, and 13,391 unique teachers.

---

<sup>7</sup> See [http://www.childandfamilypolicy.duke.edu/project\\_detail.php?id=35](http://www.childandfamilypolicy.duke.edu/project_detail.php?id=35) for additional information. See Goldhaber (2007), Jackson (2013), and Rothstein (2010) for examples of other studies that have fit VA models to the NCERDC data.

<sup>8</sup> Students were matched to teachers using administrative roster data (Course Membership file) that accurately link students and teachers to courses. Such records exist for over 80% of students. Because absences can be affected by multiple teachers, the sample is restricted to self-contained teachers who taught the student both math and reading.

Table 1 summarizes the variation in student absences and the composition of the analytic sample. The average student was absent about 6 times per year and the standard deviation (SD) of about 5.5 indicates that there is substantial variation across student-years in the sample. I decompose the variation in student absences between schools, school years, teachers, classrooms, and students by estimating corresponding “within-unit” SD in absences by computing the SD of the residuals of regressions of student absences on sets of school, school-by-year, teacher, classroom, or student fixed effects. The within-school and within-school year SD are quite similar to the overall SD, indicating that most of the variation in student absences exists within, as opposed to between, schools. The within-teacher and within-classroom SD are slightly smaller, though still constitute 95 to 97% of the variation in student absences. Again, this indicates that within schools, most variation in student absences exists within, as opposed to between, classrooms. Interestingly, the within-student SD is substantially smaller, indicating that about one third of the variation in student absences is due to within-student changes in absence rates over time. While this suggests that absences are somewhat “sticky,” there is significant within-student variation in absences over time that might be partially attributable to teachers.

#### 4. Identification Strategy

Teacher effects on student absences are identified by estimating VA models of the form:

$$A_{ijgst} = \alpha A_{i,t-1} + \beta \mathbf{x}_{it} + \gamma \mathbf{c}_{-i,jgst} + \theta_j + \pi_g + \omega_{st} + u_{ijgst}, \quad (1)$$

where  $i, j, g, s,$  and  $t$  index students, teachers, grades, schools, and years, respectively;  $A$  is annual student absences, standardized by grade and year to facilitate comparisons with the achievement results;  $\mathbf{x}$  is a vector of observed student characteristics including race, gender, poverty status, special education, and English language proficiency;  $\mathbf{c}$  is a vector of classroom

characteristics including class size, class composition, and the average of student  $i$ 's classmates' lagged absences and lagged achievement;  $\theta$ ,  $\pi$ , and  $\omega$ , are teacher, grade, and school-by-year fixed effects (FE), respectively; and  $u$  is an idiosyncratic error term.<sup>9</sup>

The school-by-year FE are central to the identification strategy and imply that the teacher effects in equation (1) are identified by comparing teachers who were in the same school during the same academic year.<sup>10</sup> Importantly, this controls for the sorting of teachers across schools, non-parametric school time trends, and variation across both schools and time in the length of academic calendars. The latter is important in the current context because longer school calendars provide more opportunities to be absent. Moreover, school-by-year FE control for school-level leadership and policy changes that either directly influence student attendance or the way that student absences are administratively recorded.

Ordinary Least Squares (OLS) is taken as the preferred estimator of equation (1) for two reasons. First, Guarino, Reckase, and Wooldridge (2015) find OLS to be the most robust estimator to a variety of potential student-teacher assignment scenarios. This is potentially important, as Rothstein (2010) finds evidence of non-random sorting in North Carolina. Second, Chetty et al. (2013) find that most sorting of students to teachers is based on lagged test scores and that conditioning on lagged test scores alone yields estimated teacher effects with near-zero

---

<sup>9</sup> Online Appendix Table A.1 investigates the sensitivity of the main results to using three alternative definitions of student absences: levels (unstandardized), natural logs, and indicators for “chronically absent.” These results are consistent with the main results reported in table 2.

<sup>10</sup> However, some policy makers may wish to compare teachers within or between schools and different specifications can produce different rankings (e.g., Goldhaber and Theobald 2012). Online Table A.2 examines the sensitivity of the main results reported in table 2 by replacing the school-by-year FE with school FE or removing them altogether. Identification in the former is driven by teachers who changed schools during the sample time period. The latter provide state-wide teacher comparisons. Both sets of estimates show qualitatively similar patterns to those generated by the preferred baseline specification of equation (1).

bias. Similarly, Kane and Staiger (2008) find that controlling for lagged test scores yields unbiased estimates of teacher effects and that controlling for average classroom characteristics (i.e., the vector  $\mathbf{c}$ ) improves the precision of estimated teacher effects. I also consider an extension of (1) that conditions on lagged test scores and lagged absences, which produces qualitatively similar estimates.

Still, even after conditioning on observed student and classroom characteristics, within school-year endogenous sorting of students to teachers remains a threat to identification. Accordingly, I follow Jackson (2013) and Chetty et al. (2013) in testing for endogenous sorting based on observable student characteristics by regressing predicted outcomes on estimated out-of-sample teacher effects and school-by-year FE in the following linear regression model:

$$\hat{y}_{ijgst} = \delta \hat{\theta}_{jt} + \omega_{st} + u_{ijgst}. \quad (2)$$

The  $\hat{\theta}$  in equation (2) are year-specific out-of-sample teacher effects estimated by equation (1) using all non- $t$  years of data. The  $\hat{y}$  in (2) are fitted values from OLS regressions of actual student absences, math scores, and reading scores on their lagged values and observed student characteristics. Intuitively, a significant correlation between teacher effectiveness and predicted student outcomes is suggestive of endogenous sorting. The estimated sign of  $\delta$  speaks to the type of sorting (e.g., a positive  $\delta$  means that on average high-performing students are assigned to more effective teachers). The results of these tests, presented in section 5, provide no evidence that OLS estimates of equation (1) are biased by endogenous sorting based on observables.

Next, I quantify the magnitude and variation in estimated teacher effects by testing their joint significance, estimating the SD of the teacher effect estimates, and by comparing estimated teacher effects at different points in the distribution (e.g., 25<sup>th</sup> versus 75<sup>th</sup> percentiles). I estimate SD of the estimated teacher effects by following the two-step procedure outlined in Jackson

(2013, p. 14), which follows from Kane and Staiger (2008). First, I compute classroom-level average residuals from estimates of equation (1) that leave the teacher effects in the model's error term. Second, I compute the covariance between each classroom's average residual and that from a randomly-chosen classroom taught by the same teacher in a different year. To avoid potentially compromising effects of outliers, I repeat step two 50 times and report the median estimated SD (Jackson 2013). This approach is preferred to estimating the SD of estimated teacher FE because it eliminates variation due to both sampling error and unobserved classroom shocks that are not associated with teacher effectiveness.

To facilitate comparisons of the magnitude and distribution of estimated teacher effects on student absences to those on academic achievement I estimate traditional VA model analogs to equation (1) that replace  $A$  with math and reading test scores. All test scores are standardized by grade, year, and subject to have mean zero and SD one (Ballou 2009). The achievement VA models also condition on current student absences, which raises a potentially interesting modeling question, though in practice models that do and do not control for current student absences produce nearly identical results.<sup>11</sup> In section 7, these estimates are used to examine the cross-domain stability of VA measures of teacher effectiveness by comparing VA-based

---

<sup>11</sup> Specifically, because absences are at least partly outside of teachers' control, it is unclear whether absences should be controlled for in VA models designed to identify teachers' effects on academic achievement (Harris 2011; Noell et al. 2008). On the one hand, at least some student absences are completely outside teachers' control and these absences should unambiguously be controlled for (Harris 2011). On the other hand, absences caused by teachers are outcomes of the education production function and are thus "bad controls" (e.g., Angrist and Pischke 2009, p. 64). In practice, the precise number of absences caused teachers is unknown and analysts are left with two suboptimal options: either omit student absences from the VA model and suffer from potential omitted variables bias or control for student absences at the risk of "over controlling" and penalizing teachers who improve test scores via improving student attendance. Online Appendix Table A.3 shows that this is a practically unimportant modeling decision, as rankings of teacher effectiveness generated by VA models that do condition on student absences are nearly identical to rankings generated by VA models that do not condition on student absences.

rankings of teachers' effects on student absences to analogous rankings of teachers' effects on academic achievement.

## 5. Main Results

### <A>Teacher Effects on Student Absences</A>

Table 2 summarizes estimates of equation (1) for fourth and fifth graders' absences, math achievement, and reading achievement. The baseline estimates of teachers' effects on student absences reported in column 1 are strongly jointly significant and exhibit significant variation across teachers: the Kane and Staiger (2008) consistent estimate of the SD of teacher effects on absences is 0.07 of an absence SD. The difference between the effect of a 90<sup>th</sup> percentile teacher and a 10<sup>th</sup> percentile teacher is about 90% of an absence SD, while the difference between teachers at the first and third quartiles is about 40% of an absence SD. Column 2 shows that the preferred baseline results reported in column 1 are robust to controlling for lagged test scores.

Columns 3 and 4 of table 2 report estimates for math and reading achievement, respectively. The estimated coefficients on student absences are negative, statistically significant, and similar in magnitude to estimates reported in the existing literature (e.g., Aucejo and Romano 2013; Gershenson et al. 2014; Gottfried 2009). Consistent with prior research on teacher effectiveness, the results reported in columns 3 and 4 of table 2 suggest that teachers have greater influence on students' math achievement than on reading achievement (e.g., Hanushek and Rivkin 2010; Jackson 2013; Kane and Staiger 2008; Rockoff 2004). Moreover, the estimated SD of teacher effects on math and reading are similar in magnitude to those found in previous studies of primary school teachers in North Carolina (e.g., Rothstein 2010) and across the U.S. (Hanushek and Rivkin 2010).

Interestingly, the estimated SD of teacher effects on student absences reported in columns 1 and 2 of table 2 are similar in magnitude to those of teacher effects on both math and reading achievement. Indeed, they are identical to those for reading. Taken together, the results reported in table 2 suggest that the total variation in teachers' effects on student absences is similar to that in teachers' effects on academic achievement.

The bottom panel of table 2 provides evidence that the estimated teacher effects are not biased by endogenous sorting of students to teachers based on observable characteristics. Specifically, none of the estimated coefficients on the out-of-sample estimates of teacher quality ( $\delta$  in equation 2) are significantly different from zero at traditional confidence levels. Moreover, the estimated coefficients and corresponding standard errors are relatively small in magnitude. This is reassuring and suggests that the teacher effect estimates from equation (1) are causal.

#### <B>External Validity of Main Results<B>

The generalizability of any state-level analysis is a concern, even in as diverse a state as North Carolina. Accordingly, I augment the main results presented above with similar analyses of the nationally representative Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K). The ECLS-K is a longitudinal dataset collected by the National Center for Education Statistics (NCES). The original sample of approximately 22,000 children from about 1,000 kindergarten programs was designed to be nationally representative of kindergartners during the 1998-99 academic year. Subsequent analyses of the ECLS-K data are weighted by sampling weights provided by NCES that adjust for the oversampling of certain demographic groups.<sup>12</sup> Importantly, the ECLS-K administered age-appropriate math and reading assessments each

---

<sup>12</sup> Specifically, I use the C#CW0 longitudinal weight, where # is wave number.

spring and asked school administrators to report each student’s total annual absences.

The ECLS-K surveyed children, parents, teachers, and school administrators during the fall and spring of kindergarten and the spring of first, third, and fifth grades. As a result, VA models similar to equation (1) can only be estimated for first-grade students conditional on kindergarten absences. Like in the North Carolina analysis, students who experienced a mid-year classroom change, repeated kindergarten or first grade, or are missing test-score or demographic data are excluded from the analysis. The analytic sample is also restricted to classrooms in which at least five students were sampled by the ECLS-K, so that there are a reasonable number of data points with which to estimate classroom effects. These exclusions yield an analytic sample of 2,350 student-year observations.<sup>13</sup> The reference to classrooms and not teachers is intentional, as the ECLS-K followed one cohort of students and observes each teacher in only one year.

Specifically, the ECLS-K analog to equation (1) is

$$A_{ijgst} = \alpha A_{i,t-1} + \beta \mathbf{x}_{it} + \lambda_{jgst} + u_{ijgst}, \quad (3)$$

where  $\lambda$  is a classroom FE. Importantly, the classroom effects in (3) can neither be interpreted as, nor decomposed into, teacher effects. For example, the classroom FE specification of (3) cannot distinguish teacher effects from class size effects, as the classroom effects are treated as fixed rather than random and classrooms are nested within teachers. As a result, the ECLS-K results cannot be directly compared to the analyses of North Carolina teacher effects discussed above. Given that the ECLS-K follows one cohort of students over time, a teacher FE specification equivalent to equation (1) cannot be estimated using the ECLS-K data because each teacher is only observed in one academic year. Similarly, the school, grade, and year FE commonly included in VA models (e.g., equation 1) are subsumed by the classroom FE in (3). Nonetheless,

---

<sup>13</sup> Reported ECLS-K sample sizes are rounded to the nearest 50. See Gershenson et al. (2014) for further discussion of the ECLS-K’s student assessments and absence data.

equation (3) can be estimated using both the North Carolina and ECLS-K data. The generalizability of the main results can then be inferred by comparing estimates of (3) using the North Carolina data to estimates of (3) using the ECLS-K data.

Table 3 summarizes the variation in estimated classroom effects in both datasets. The similarities across datasets are striking. For example, the differences between classroom effects at the 25<sup>th</sup> and 75<sup>th</sup> percentiles are about one third of a standard deviation for each outcome in each dataset. Together, the results reported in table 3 suggest that the analysis of North Carolina teachers is at least somewhat representative of public primary school teachers in the U.S.

## **6. Further Analyses of Teachers' Effects on Student Absences**

Having shown that the main results are arguably internally and externally valid, this section provides three additional pieces of evidence regarding the internal validity of the finding that teachers affect student absences. Specifically, this section investigates the extent to which individual teachers' effects on student absences are stable over time, the persistence of grade- $g$  teachers' effects on student absences in grade  $g+1$ , and whether the ability to improve student attendance evolves over teachers' careers. In doing so, the results presented in this section shed some light on the mechanisms through which teachers affect students' attendance. Taken as a whole, these results lend additional empirical support to the general finding that teachers modestly affect student attendance.

### **<C>Intertemporal Stability of Estimated Teacher Effects on Student Absences<C>**

If the teacher effects discussed in section 5 merely reflect noise or the composition of teachers' classrooms in specific years, then the intertemporal stability of teachers'

contemporaneous effects on student absences would be indistinguishable from zero. Alternatively, if there is a stable component in teachers' ability to influence student absences, teacher rankings should be positively correlated across years. Accordingly, table 4 reports two types of intertemporal Spearman Rank Correlations of teachers' effects on student absences, math achievement, and reading achievement. The top panel of table 4 compares teacher rankings generated by data from the 2006-07 and 2007-08 school years to teacher rankings generated by data from the 2008-09 and 2009-10 school years. These teacher effects were generated by estimating equation (1) separately for each of the two two-year time periods for teachers for whom data are available for all four years. The bottom panel of table 4 compares classroom rankings across each pair of contiguous years and the weighted average of these three correlations. These classroom effects were generated by estimating equation (3) separately for each year and comparing the resulting rankings for teachers who taught in two consecutive years. Both panels of table 4 provide evidence that is consistent with significant teacher effects on student absences, as the intertemporal rank correlations are about one tenth of a SD and are strongly statistically significant. Nonetheless, the intertemporal rank correlations in teachers' effects on student absences are only about one fourth to one half the size of the intertemporal rank correlations in teachers' effects on test scores. This could be because estimated effects on student absences are noisier or because teachers' abilities to affect student absences are more limited and context-dependent than their abilities to affect test scores. Finally, estimates of the intertemporal stability of teachers' effects on test scores are consistent with those in the existing literature (e.g., Goldhaber and Hansen 2013; Loeb and Candelaria 2012; McCaffrey et al. 2009).

<D>Persistence of Teachers' Effects on Student Absences<D>

The mechanisms through which teachers can affect student absences, discussed in section 2, suggest that teachers' effects on student attendance should persist in subsequent school years. For example, a teacher who instills a love of learning in students or who successfully motivates parents to facilitate regular attendance will likely affect students' current and future attendance. To test whether this is the case, I use the method proposed by Jacob, Lefgren, and Sims (2010) to estimate the average persistence of fourth grade teachers' effects on students' fifth-grade outcomes. Specifically, Jacob et al. (2010) show that the OLS estimate of  $\alpha$  in equation (3) can be interpreted as the persistence of observed outcome  $y$  ( $\alpha^{\text{OLS}}$ ), while the instrumental variables (IV) estimate of  $\alpha$  that instruments for  $y_{i,t-1}$  with  $y_{i,t-2}$  can be interpreted as the persistence of the long-run component of  $y$  ( $\alpha^{\text{LR}}$ ). Finally, the authors show that the IV estimate of  $\alpha$  that instead instruments for  $y_{i,t-1}$  with  $\hat{\theta}_{i,t-1}$ , as defined in equation (2), can be interpreted as the fraction of variation in the LR component of  $y$  attributable to teachers. Accordingly, the third estimate of  $\alpha$  represents the average persistence of teacher effects ( $\alpha^{\text{P}}$ ).

Table 5 reports each of these three estimates of  $\alpha$  for absences, math achievement, and reading achievement. Estimates of  $\alpha^{\text{OLS}}$  and  $\alpha^{\text{LR}}$  are slightly smaller for absences than for math and reading achievement, suggesting that absences are less persistent over time than academic skills. This is consistent with the general result that non-cognitive skills are more malleable than cognitive skills (e.g., Cunha and Heckman 2008; Heckman 2000). It is also reassuring that the estimates of  $\alpha^{\text{OLS}}$  and  $\alpha^{\text{LR}}$  for math and reading reported in table 5 are similar to the corresponding estimates reported by Jacob et al. (2010). Interestingly, the estimate of  $\alpha^{\text{P}}$  for absences is larger than the corresponding estimates for math and reading achievement, though the absence estimate is less precisely estimated. Still, the null hypothesis of zero persistence in teachers' effects on absences can be rejected with 5% confidence. This suggests that teachers'

effects on student absences are at least as persistent as teachers' effects on academic achievement, despite less intertemporal persistence in the LR component of students' absences than in the LR components of math and reading ability. Specifically, the point estimate of 0.51 reported in column 1 of table 5 suggests that about half the variation in fourth-grade student absences attributable to fourth-grade teachers persists in fifth grade. Again, this result is consistent with the general finding that teachers affect student attendance.

#### <E>Does Teaching Experience Affect Student Absences?<E>

Finally, if teachers do affect student attendance, it stands to reason that their ability to do so improves with teaching experience (Ladd and Sorensen 2014). For example, more experienced teachers might converse with parents and teach character skills more effectively than their less-experienced counterparts. Evidence of an “experience gradient” in teachers' effects on student attendance would lend additional empirical support to the claim that teachers affect student attendance. Accordingly, I estimate the effect of teachers' experience on student absences using the non-parametric specification and estimation framework advocated by Wiswall (2013). Specifically, the returns to teaching experience are estimated in a two-step procedure. First, the classroom fixed effects in equation (3) are estimated ( $\hat{\lambda}$ ) and saved for use in step 2. Second, the  $\lambda$  are regressed on teacher experience (*exper*); teacher, grade, and year FE; and the vector of classroom characteristics from equation (1). Following Wiswall (2013), I model teachers' experience as a set of  $K = 36$  binary indicators for each experience level from 1 to 35 plus a category for 36+ years of experience, where new teachers with zero experience comprise the omitted reference category. Formally,

$$\hat{\lambda}_{ijgst} = \sum_{k=1}^K \varphi_k 1\{exper_{jt} = k\} + \gamma \mathbf{c}_{-i,jgst} + \theta_j + \pi_g + \delta_t + u_{ijgst}, \quad (4)$$

where  $1\{\cdot\}$  is the indicator function.<sup>14</sup>

The 36 estimated  $\varphi$  parameters for each outcome (i.e., absences, math achievement, and reading achievement) are plotted in figure 2. For math, the nonparametric estimates suggest that returns to experience continue to accrue over the first twenty years of teaching, which are consistent with the findings of Wiswall (2013). The relationship between teaching experience and students' reading achievement follows a similar pattern, but the effects are only about half as large as those for math achievement. Again, this is consistent with previous research on the returns to teaching experience (e.g., Clotfelter et al. 2007; Kane, Rockoff, and Staiger 2008; Ladd and Sorensen 2014) and with results presented earlier in this article that suggest that teacher effects on reading achievement are about half as large as those on math achievement (e.g., table 2).

The estimated relationship between teaching experience and student absences mirrors that between teaching experience and reading achievement, suggesting that more experienced teachers are modestly more effective at reducing student absences. For example, on average, students assigned to teachers who have 20 years of teaching experience have about 20% of a SD fewer absences than similar students assigned to new teachers. Generally, these effects are smaller in magnitude than the effects of middle school math and English teachers' on middle school student absences in North Carolina found by Ladd and Sorensen (2014). For example, the authors find that teachers who have 20 years of experience decrease student absences by about 60% of a student-absence SD. This difference could result from middle school students having

---

<sup>14</sup> School FE are omitted from equation (4) because relatively few teachers in the analytic sample changed schools.

relatively more agency over their absences than primary school students. Still, that the effects of primary school teachers' teaching experience on student absences shown in figure 2 are similar in magnitude to those on reading achievement is again consistent with the main results presented in table 2 and suggestive of a causal relationship between teacher effectiveness and student attendance.

## **7. Cross-Domain Stability of Estimated Teacher Effects**

I now compare the estimated teacher effects on student absences to those on academic achievement to examine the stability of teacher effectiveness across cognitive and non-cognitive domains. I do so by comparing rankings of teacher and classroom effectiveness based on the teacher and classroom effects generated by equations (1) and (3). Comparisons are made between rankings rather than between point estimates because VA models frequently produce reliable rankings of teacher effectiveness even when the point estimates are inconsistent (Guarino, Reckase, Stacy, and Wooldridge 2014) and rankings are arguably more policy relevant than point estimates. Specifically, rankings are compared across domains in three ways. First, I compute Spearman rank correlations. Second, I compute the percentage of teachers who are above average in both rankings, and similarly for various quantiles of interest. Finally, more nuanced transition matrixes are reported in Online Appendix Table A.4.

The first panel of table 6 summarizes these relationships for teachers. Spearman rank correlations between the absence and academic achievement rankings are close to zero and actually negative, suggesting that teachers who excel in one domain do not necessarily excel along others. This is further evidenced by the fact that relatively few teachers are above or below specific thresholds in both rankings. For example, only about 1 to 2% of teachers are in the top

(bottom) decile, 7% are in the top (bottom) quartile, and 25% are in the top (bottom) half of both the absence and academic rankings.<sup>15</sup> There is noticeably more stability between the math and reading rankings. It is reassuring that the cross-subject correlation of 0.34 fits comfortably within the range of previous estimates (e.g., Goldhaber et al. 2013; Loeb and Candelaria 2012).

The second and third panels of table 6 compare the cross-domain and cross-subject stability of estimated classroom effects using the NCERDC and ECLS-K datasets, respectively. Once again, the North Carolina and ECLS-K analyses yield remarkably similar results, suggesting that the North Carolina results generalize to the U.S. population. The cross-domain rank correlations are close to zero in both datasets and the cross-subject rank correlations are about 0.4. Taken together, the results presented in table 6 suggest that teachers who are (in)effective in one domain are not necessarily (in)effective in others. This result is consistent with research by Jackson (2013) and Jennings and DiPrete (2010) and suggests that narrowly focusing on test scores will potentially misclassify teachers who improve students' character skills such as regular attendance as ineffective.

## **8. Discussion**

This paper uses longitudinal administrative data on teachers and students in North Carolina to estimate teacher effects on both student absences and academic achievement. The

---

<sup>15</sup> These numbers can be converted into the percentage of eligible teachers by dividing by the quantile's range. For example, 17% ( $1.7/10$ ) of teachers in the top decile of the absence ranking are in the top decile of the math ranking. It is also worth noting that table A.4 in the online appendix reports transition matrixes that provide a more nuanced view of the cross-domain and cross-subject stability of teacher rankings, as correlations can mask large swings in rankings (Goldhaber and Theobald 2012). Again, rankings are less stable across domains than across subjects. For example, only about 22% of teachers in the top (bottom) fifth of the math rankings are also in the top (bottom) fifth of the absence rankings. The cross-subject transition matrixes are consistent with previous research that finds about 40% of the lowest (highest) performing teachers in math are similarly low (high) performing in reading (e.g., Loeb et al. 2012).

analyses yield two novel findings, which are generally consistent with similar analyses of the nationally representative ECLS-K and robust to a variety of VA model specifications. First, teachers have statistically significant effects on student absences, which are not biased by endogenous sorting of students to teachers based on observable student characteristics, and are similar in magnitude to teachers' effects on reading test scores. Second, there is essentially zero correlation between rankings of teacher effects on absences and rankings of teacher effects on academic achievement, which suggests that there are multiple dimensions of effective teaching and that teachers who excel along one dimension do not necessarily excel along others. These findings are generally consistent with previous studies of teachers' ability to affect non-cognitive and socio-behavioral skills in other contexts (Jackson 2013; Jennings and DiPrete 2010).

Three additional results lend additional support to the finding that teachers affect student absences. First, teachers' contemporaneous effects on student absences are positively correlated over time, suggesting that there is a permanent component to teachers' effects on student absences over and above transitory components associated with a particular classroom. Second, teachers' effects on student absences persist into the following academic year, suggesting that teachers affect students' (or parents') attitudes and preferences, rather than simply providing short-run incentives to attend class. Third, more experienced teachers tend to have larger effects on student attendance. This could either be because learning to alter student behaviors such as attendance takes time or because new teachers initially choose to focus on improving their academic instructional skills. It would be useful for future research to investigate the underlying sources of the experience gradient in teachers' effects on character skills. Similarly, future research might extend the analyses conducted in section 6 to other contexts and investigate how teachers affect other types of character skills and related student outcomes.

The results presented in this article contribute to two distinct literatures in the economics of education, as well as to our understanding of the educational process more generally. First, the finding that teachers affect primary school students' attendance furthers our understanding of the education production function and the educational inputs that develop character skills. Based on existing estimates of student absences' effects on test scores (e.g., Aucejo and Romano 2013; Gershenson et al. 2014), the decrease in student absences attributable to a one SD improvement in teacher effectiveness (i.e., 0.07 absence SD, or about 0.4 student absences) translates into relatively small achievement gains comparable to about 3% of the test-score gains attributable to a one SD improvement in teacher effectiveness (as measured by effects on test scores).

However, student absences, particularly in primary school, are important over and above their direct impact on test scores for several reasons. Children form habits and undergo substantial developmental changes during these formative years, which is important given that high school absences predict negative long-run outcomes such as grade retention (Nield and Balfanz 2006), drug use (Hallfors et al. 2002), and dropping out of school (Rumberger and Thomas 2000) and the longer-term importance of regular attendance in the labor market (Heckman and Kautz 2013). Indeed, improving attendance habits could be one mechanism through which primary school teachers affect long-run socioeconomic outcomes. Moreover, there are positive externalities, or peer effects, of individual students' attendance and school engagement on the academic achievement of their classmates (Gottfried 2011). It could also be the case that student absences influence peers' attendance habits. Finally, from a practical standpoint, the finding that teachers affect student attendance is likely to be of direct interest to school principals and administrators seeking to improve the academic performance and school engagement of disadvantaged and chronically absent students. Information on the teachers who

most effectively improve student attendance might be used either to identify the classroom characteristics and teaching strategies that contribute to improvements in student attendance or to strategically assign students to teachers. Similarly, the finding that teachers can and do affect student attendance might be used in pre-service and professional-development training to motivate the importance and ability of teachers to influence students' socio-emotional behaviors.

Second, the current study also contributes to the general literature on the use and estimation of VA models of teacher effectiveness. The small and sometimes negative correlation between rankings of teacher effectiveness across domains (absences vs. academic achievement) suggests the importance of evaluating teachers along multiple objective dimensions. The lack of a strong positive relationship between these rankings could result from some teachers eliciting test score gains by running a strict, drill-based classroom at the expense of maintaining a stimulating learning environment. Alternatively, if teachers who improve attendance have larger average class sizes as a result, a "bad-apple" model of peer effects (Lazear 2001) might undermine the classroom's average test performance. It would be useful for future research to further investigate the within-teacher relationships between different types of teaching skills. Similarly, future research might probe the exact mechanisms through which teachers affect attendance and related character skills, as the current study is unable to disentangle effects on parental involvement from effects on students' dispositions and behaviors. In the meantime, however, current teacher evaluation systems that prioritize teachers' effects on student test scores are potentially failing to recognize the effectiveness of teachers who facilitate students' development in other domains and potentially divert teachers' time and energy away from lessons and activities that develop character skills. It is therefore important that teacher evaluation systems include multiple measures of teacher effectiveness, perhaps including

measures of teachers' ability to improve students' attendance and related character skills. Indeed, if teachers were incentivized and encouraged to improve student attendance and related character skills, their effects on such behaviors would likely be even larger than those found here.

## References

- Alexander, K., D. Entwisle, and N. Kabbani. 2001. The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record*, 103(5): 760-822.
- Almlund, M., A. L. Duckworth, J. J. Heckman, and T. Kautz. 2012. Personality psychology and economics. In *Handbook of economics of education Vol. 4*, edited by E.A. Hanushek, S.J. Machin, and L. Woessmann. Amsterdam: North Holland.
- Angrist, J., and S. Pischke. 2009. *Mostly harmless econometrics: An empiricists' companion*. Princeton, NJ: Princeton Univ. Press.
- Aucejo, E. M., and T. F. Romano. 2013. Assessing the effect of school days and absences on test score performance. Unpublished paper, Duke University.
- Baker, E. L., P. E. Barton, L. Darling-Hammond, E. Haertel, H. F. Ladd, R. L. Linn, and L. A. Shepard. 2010. *Problems with the use of student test scores to evaluate teachers* Vol. 278. Washington, DC: Economic Policy Institute.
- Ballou, D. 2009. Test scaling and value-added measurement. *Education Finance and Policy* 4(4): 351-383.
- Borghans, L., A.L. Duckworth, J.J. Heckman, and B. Ter Weel. 2008. The economics and psychology of personality traits. *Journal of Human Resources* 43(4): 972-1059.
- Bruner, C., A. Discher, and H. Chang. 2011. *Chronic elementary absenteeism: A problem hidden in plain sight*. <http://www.edweek.org/media/chronicabsence-15chang.pdf>. Accessed 20 January 2014.
- Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan. 2011. How does your kindergarten classroom affect your earnings? Evidence from Project Star. *The Quarterly Journal of Economics* 126(4): 1593-1660.

Chetty, R., J.N. Friedman, and J.E. Rockoff. 2013. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. NBER Working Paper No. w19423.

Clotfelter, C. T., Ladd, H. F., and J. L. Vigdor. 2007. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review* 26(6): 673-682.

Cunha, F., and J.J. Heckman. 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43(4): 738-782.

Cunha, F., J.J. Heckman, and S.M. Schennach. 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3): 883-931.

Dobbie, W. 2011. Teacher characteristics and student achievement: Evidence from Teach for America. Unpublished paper, Princeton University.

Dombkowski, K. 2001. Will the real kindergarten please stand up? Defining and redefining the 20th century U.S. kindergarten. *History of Education* 30(6):527-45.

Duckworth, A. L., C. Peterson, M.D. Matthews, and D.R. Kelly. 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology* 92(6): 1087.

Duncan, G. J., and K. Magnuson. 2011. The nature and impact of early achievement skills, attention skills, and behavior problems. In *Whither Opportunity*, edited by G. Duncan and R. Murnane, pp. 47-69. New York: Russel Sage Foundation.

Fenstermacher, G. D., and V. Richardson. 2005. On making determinants of quality in teaching. *Teachers College Record* 107: 186-213.

Gallup Educational Polls. 2004. Available <http://www.gallup.com/poll/1612/education.aspx>. Accessed 15 April 2014.

Gershenson, S., A. Jacknowitz, and A. Brannegan. 2014. Are student absences worth the worry in U.S. primary schools? Unpublished paper, American University.

Goldhaber, D. 2007. Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources* 42(4): 765-794.

Goldhaber, D., J. Cowan, and J. Walch. 2013. Is a good elementary teacher always good? Assessing teacher performance estimates across subjects. *Economics of Education Review* 36: 216-228.

Goldhaber, D. and M. Hansen. 2013. Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica* 80(319): 589–612.

Goldhaber, D., and R. Theobald. 2012. *Do different value-added models tell us the same things?* Available <http://www.carnegieknowledge.org/briefs/value-added/different-growth-models/>. Accessed 11 November 2014.

Gottfried, M.A. 2009. Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis* 31(4): 392-419.

Gottfried, M. A. 2011. Absent peers in elementary years: The negative classroom effects of unexcused absences on standardized testing outcomes. *Teachers College Record* 113(8): 1597-1632.

Guarino, C. M., M. D. Reckase, and J. M. Wooldridge. 2015. Can value-added measures of teacher performance be trusted? *Education Finance and Policy*. In Press.

Guarino, C. M., M. D. Reckase, B. W. Stacy, and J. M. Wooldridge. 2014. Evaluating specification tests in the context of value-added estimation. IZA Discussion Paper 7974.

Hallfors, D., J. L. Vevea, B. Iritani, H. Cho, S. Khatapoush, and L. Saxe. 2002. Truancy, grade point average, and sexual activity: A meta-analysis of risk indicators for youth substance

use. *The Journal of School Health* 72(5): 205-211.

Hanushek, E. A., and S.G. Rivkin. 2010. Generalizations about using value-added measures of teacher quality. *American Economic Review* 100(2): 267-271.

Harris, D. N., 2011. *Value-added measures in education*. Cambridge, MA: Harvard Education Press.

Heckman, J. J. 2000. Policies to foster human capital. *Research in Economics* 54(1): 3-56.

Heckman, J. J., and T. Kautz. 2013. Fostering and measuring skills: Interventions that improve character and cognition. NBER Working Paper No. w19656.

Heckman, J. J., J. Stixrud, and S. Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24(3): 411-482.

Jackson, C. K., 2013. Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. NBER Working Paper No. w18624.

Jacob, B. A., L. Lefgren, and D. P. Sims. 2010. The persistence of teacher-induced learning. *Journal of Human Resources* 45(4): 915-943.

Jennings, J. L., and T. A. DiPrete. 2010. Teacher effects on cultural capital development in elementary school. *Sociology of Education* 83(2):135–159.

Kane, T. J., and D.O. Staiger. 2008. Estimating teacher impacts on student achievement: An experimental evaluation. NBER Working Paper No. w14607.

Kane, T. J., Rockoff, J. E., and D. O. Staiger. 2008. What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review* 27(6): 615-631.

Kelly, S. 2012. Understanding teacher effects: Market versus process models of educational improvement. In *Assessing Teacher Quality*, edited by S. Kelly, pp. 7-32. New York: Teachers College Press.

Koedel, C., and J.R. Betts. 2007. Re-examining the role of teacher quality in the educational production function. Unpublished paper, Vanderbilt University.

Ladd, H. F., and Sorensen, L. C. (2014) Returns to teacher experience: Student achievement and motivation in middle school. Calder Working Paper No. 112.

Lerman, R. I. 2013. Are employability skills learned in US youth education and training programs? *IZA Journal of Labor Policy* 2(1): 6.

Lockwood, J. R., D.F. McCaffrey, L.S. Hamilton, B. Stecher, V.N. Le, and J.F. Martinez. 2007. The Sensitivity of Value-Added Teacher Effect Estimates to Different Mathematics Achievement Measures. *Journal of Educational Measurement* 44(1): 47-67.

Loeb, S., and C. Candelaria. 2012. How stable are value-added estimates across years, subjects, and student groups? Available <http://www.carnegieknowledgenetwork.org/briefs/value-added/value-added-stability/>. Accessed 11 November 2014.

Loeb, S., D. Kalogrides, and T. Béteille. 2012. Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy* 7(3): 269-304.

Lounsbury, J. W., R.P. Steel, J.M. Loveland, and L.W. Gibson. 2004. An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence* 33(5): 457-466.

Lundberg, S. 2012. Personality and marital surplus. *IZA Journal of Labor Economics* 1(1): 1-21.

Lundberg, S. 2013. The college type: Personality and educational inequality. *Journal of Labor Economics* 31(3): 421-441.

McCaffrey, D. F., T.R. Sass, J.R. Lockwood, and K. Mihaly. 2009. The intertemporal variability of teacher effect estimates. *Education Finance and Policy* 4(4): 572-606.

- Monk, D. H., and M.A. Ibrahim. 1984. Patterns of absence and pupil achievement. *American Educational Research Journal* 21(2): 295-310.
- Morrison, T., B. Maciejewski, C. Giffi, E.S. DeRocco, J. McNelly, and G. Carrick. 2011. *Boiling point? The skills gap in US manufacturing*. Washington DC: The Manufacturing Institute.
- Neild, R. C., and R. Balfanz. 2006. An extreme degree of difficulty: The educational demographics of urban neighborhood high schools. *Journal of Education for Students Placed at Risk* 11(2): 123-141.
- Noell, G. H., B. A. Porter, R. M. Patt, and A. Dahir. 2008. *Value added assessment of teacher preparation in Louisiana: 2004-2005 to 2006-2007*. Unpublished paper, Louisiana State University.
- Pritchard, J. 2013. *The Importance of Soft Skills in Entry-Level Employment and Postsecondary Success: Perspectives from Employers and Community Colleges*. Seattle, WA: Seattle Jobs Initiative
- Rivkin, S. G., E.A. Hanushek, and J.F. Kain. 2005. Teachers, schools, and academic achievement. *Econometrica* 73(2): 417-458.
- Rockoff, J. E. 2004. The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review* 94(2): 247-252.
- Rothstein, J. 2010. Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics* 125(1): 175-214.
- Rumberger, R. W. 1995. Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal* 32(3): 583-625.
- Wiswall, M. 2013. The dynamics of teacher quality. *Journal of Public Economics* 100: 61-78.

**Table 1: North Carolina Analytic Sample Summary Statistics**

Variable	Mean	SD
Absences	6.10	(5.52)
(within school)		(5.44)
(within school-year)		(5.40)
(within teacher)		(5.34)
(within classroom)		(5.22)
(within student)		(1.75)
<hr/> Standardized (Mean 0, SD 1)		
Absences	-0.02	(0.96)
Math score	0.09	(0.97)
Reading score	0.07	(0.97)
<hr/> Lagged standardized		
Absences	-0.04	(0.96)
Math score	0.10	(0.96)
Reading score	0.09	(0.95)
Fourth grade	0.56	
Fifth grade	0.44	
<hr/> Child race/ethnicity		
Non-Hispanic White	0.56	
Non-Hispanic Black	0.26	
Hispanic	0.11	
Other	0.07	
Female	0.51	
Below poverty level	0.46	
Limited English Proficiency (LEP)	0.01	
Math Disability	0.01	
Reading Disability	0.03	
<hr/> School year		
2006-07	0.25	
2007-08	0.25	
2008-09	0.24	
2009-10	0.26	
N (Teachers)		13,391
N (Classrooms)		27,943
N (Student Years)		446,244

Notes: SD = standard deviation. Standardized absence and test score means and SD are not precisely 0 and 1 because the standardization was performed using all available absence and test score data.

**Table 2: Baseline Teacher Effect Estimates**

Outcome:	Absences		Math	Reading
	1	2	3	4
Lagged absences	0.580 (0.003)***	0.577 (0.003)***		
Lagged math		-0.041 (0.002)***	0.780 (0.001)***	
Lagged reading		0.011 (0.002)***		0.758 (0.002)***
Current absences			-0.007 (0.000)***	-0.004 (0.000)***
Controls	Yes	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes	Yes
Sch-by-Yr FE	Yes	Yes	Yes	Yes
Adj. R <sup>2</sup>	0.38	0.38	0.73	0.68
<b>Teacher FE</b>				
Joint sig. (F statistic)	1.33***	1.33***	4.04***	1.78***
Mean	0.01	0.01	-0.02	-0.01
SD of FE	0.48	0.48	0.41	0.41
SD (K-S)	0.07	0.07	0.13	0.07
90th – 10th percentile	-0.91***	-0.91***	0.90***	0.84***
75th – 25th percentile	-0.39***	-0.40***	0.43***	0.38***
<b>Sorting Test (N = 380,670)</b>				
	Predicted Absences		Predicted Math	Predicted Reading
$\hat{\theta}$ (Eq. 2)	0.002 (0.002)	0.002 (0.002)	0.002 (0.005)	0.002 (0.005)

Notes: N = 446,244 student-year observations taught by 13,391 unique teachers. Standard errors are clustered by school. Controls include indicators of child's race/ethnicity, poverty status, limited English proficiency, administratively classified learning disability, year indicators, and classroom characteristics including class size, lagged peer achievement and absences, percent of classroom eligible for free or reduced price lunch, and classroom racial composition. Absences and test scores are standardized by subject, grade, and year to have mean zero and standard deviation (SD) one. K-S refers to Kane and Staiger's (2008) method for computing consistent estimates of the SD of estimated teacher effects. The sorting test is described by equation (2) in the text.

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

**Table 3: Classroom Effect Estimates in North Carolina and the ECLS-K**

Outcome:	Absences		Math	Reading
	1	2	3	4
Lagged absences	Yes	Yes	No	No
Lagged math	No	Yes	Yes	No
Lagged reading	No	Yes	No	Yes
Current absences	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Classroom FE	Yes	Yes	Yes	Yes

**North Carolina**

Joint sig. (F stat)	1.68***	1.68***	4.16***	1.98***
Mean	0.01	0.01	-0.01	-0.01
SD of FE	0.33	0.33	0.28	0.23
90th – 10th percentile	-0.62***	-0.62***	0.68***	0.52***
75th – 25th percentile	-0.31***	-0.31***	0.35***	0.27***

**ECLS-K**

Joint sig. (F)	76.6***	97.5***	80.9***	209.8***
Mean	0.0003	0.001	0.22	0.08
SD of FE	0.39	0.39	0.28	0.27
90th – 10th percentile	-0.53***	-0.53***	0.69***	0.66***
75th – 25th percentile	-0.27***	-0.27***	0.38***	0.34***

Notes: The North Carolina sample contains 446,244 student-year observations and 27,943 classrooms. The ECLS-K sample contains 2,350 first grade students and 300 classrooms (sample sizes rounded to nearest 50). Standard errors are clustered by school. Student controls include indicators of mother's educational attainment, child's race/ethnicity, poverty status, English spoken at home, and special education designation. Absences and test scores are standardized by subject, grade, and year to have mean zero and standard deviation (SD) one. The four classroom-FE specifications reported here correspond to the four teacher-FE specifications reported in table 2.

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

**Table 4: Intertemporal Stability of Teacher Effect Estimates**

Outcome:	Absences	Math	Reading	N (teachers)
<u>Two-year teacher effects</u>				
2007-2008 to 2009-2010	0.13***	0.45***	0.23***	2,250
<u>Classroom effects</u>				
2007 to 2008	0.05***	0.43***	0.23***	4,557
2008 to 2009	0.09***	0.39***	0.24***	4,439
2009 to 2010	0.10***	0.39***	0.22***	4,571
Weighted average	0.08***	0.40***	0.23***	.

Notes: Spearman Rank Correlations are reported. Two-year teacher effects are estimated by splitting the data in two two-year samples and estimating equation (1) twice: once using 2006-07 and 2007-08 data and once using 2008-09 and 2009-10 data. Year-specific classroom effects come from estimating equation (3) separately for each school year between 2006-07 and 2009-10. The absence, math, and reading estimates are based on the preferred specifications reported in columns 1, 3, and 4 of tables 2 and 3. \*\*\* p<0.01.

**Table 5: One-Year Persistence of Fourth-Grade Teacher Effects**

Outcome:	Absences 1	Math 2	Reading 3
OLS	0.59 (0.005)***	0.78 (0.002)***	0.75 (0.002)***
Long Run (LR)	0.88 (0.005)***	0.98 (0.003)***	0.98 (0.003)***
Persistence (P)	0.51 (0.25)**	0.35 (0.06)***	0.48 (0.09)***
Classroom FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

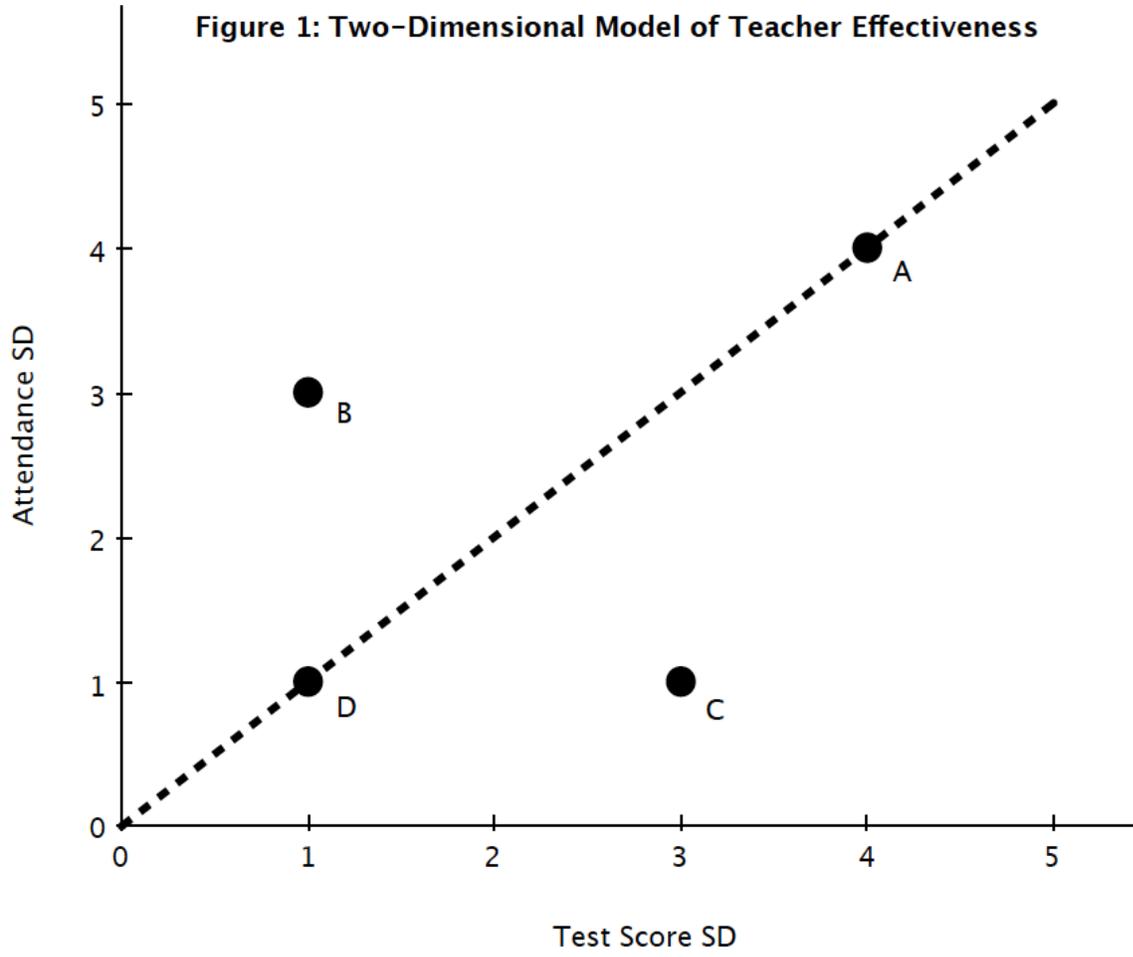
Notes: N = 101,679 fifth-grade students for whom twice-lagged test scores and absences and once-lagged out-of-sample estimated teacher quality are observed. Each cell represents the estimated coefficient on the lagged dependent variable in equation (3) from a separate regression, as described in Jacob et al. (2010). Standard errors are clustered by classroom (Jacob et al., 2010). Controls include indicators of child's race/ethnicity, poverty status, limited English proficiency, and administratively classified learning disabilities. Absences and test scores are standardized by subject, grade, and year to have mean zero and standard deviation one.

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

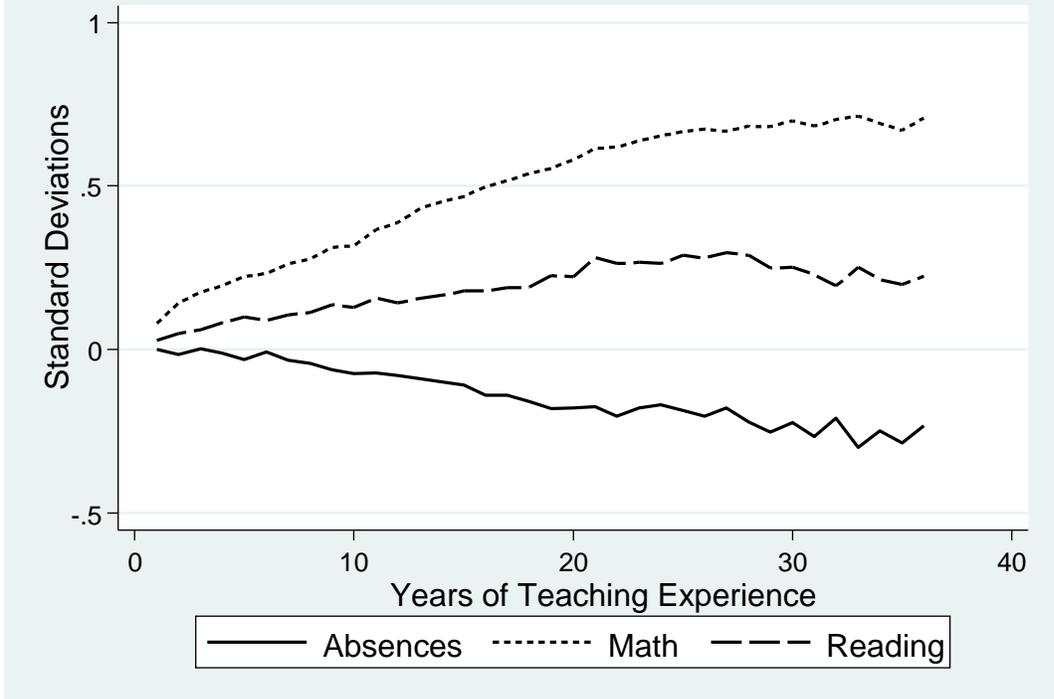
**Table 6: Cross-Domain Stability of Estimated Teacher & Classroom Rankings**

	Spearman Corr. Coeff.	Both above 90th percentile	Both above 75th percentile	Both above mean	Both above median	Both below 25th percentile	Both below 10th percentile
<b>North Carolina Teacher Rankings</b>							
Absence -Math	-0.04***	1.7%	6.5%	25.5%	24.9%	6.6%	1.1%
Absence -Reading	-0.02*	1.8%	7.2%	24.7%	25.1%	7.1%	1.6%
Math- Reading	0.34***	3.5%	11.3%	30.4%	31.0%	10.9%	2.9%
<b>North Carolina Classroom Rankings</b>							
Absence -Math	0.06***	1.3%	6.7%	27.9%	26.1%	7.1%	1.4%
Absence -Reading	0.05***	1.3%	6.7%	28.0%	25.9%	7.3%	1.7%
Math- Reading	0.46***	3.1%	11.8%	33.9%	33.2%	12.0%	3.4%
<b>ECLS-K Classroom Rankings</b>							
Absence -Math	-0.07	1.0%	5.6%	24.8%	22.9%	4.6%	0.7%
Absence -Reading	0.10*	2.0%	8.5%	28.1%	27.5%	5.9%	1.3%
Math- Reading	0.37***	2.6%	11.8%	31.7%	30.1%	9.8%	2.6%

Notes: The North Carolina sample contains 446,244 student-year observations and 27,943 classrooms. The ECLS-K sample contains 2,350 first grade students and 300 classrooms (sample sizes rounded to nearest 50). Teacher rankings are based on the teacher effects estimated in columns 1, 3, and 4 of table 2. Classroom rankings are based on the classroom effects reported in columns 1, 3, and 4 of table 3. Absence-subject and math-reading refer to cross-domain and cross-subject stability, respectively. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



**Figure 2: Nonparametric Estimates of Returns to Teaching**



Notes: This figure reports nonparametric estimates of the effect of teaching experience on three outcomes: student absences, math achievement, and reading achievement. Each outcome is standardized to have mean zero and standard deviation one. New teachers with zero years of experience comprise the omitted category, so each coefficient compares the returns to a given level of experience to a new teacher. Teachers with more than 35 years of experience are grouped together in the final category. Estimates were generated by the two-stage estimation strategy described in the text and follow from Wiswall (2013). All models control for lagged student outcomes, student characteristics, classroom characteristics, and teacher, grade, and year fixed effects. Models were estimated using the subset of NCERDC data for which teacher experience data was available. N = 295,305 student years and 9,064 unique teachers.