Engaging Teachers: Measuring the Impact of Teachers on Student Attendance in Secondary School

Jing Liu
Brown University

Susanna Loeb
Brown University

ABSTRACT: Teachers' impact on student long-run success is only partially explained by their contributions to students' short-run academic performance. For this study, we explore a second dimension of teacher effectiveness by creating measures of teachers' contributions to student class-attendance. We find systematic variation in teacher effectiveness at reducing unexcused class absences at the middle and high school level. These differences across teachers are as stable as those for student achievement, but teacher effectiveness on attendance only weakly correlates with their effects on achievement. We link these measures of teacher effectiveness to students' long-run outcomes. A high value-added to attendance teacher has a stronger impact on students' likelihood of finishing high school than does a high value-added to achievement teacher. Moreover, high value-added to attendance teachers can motivate students to pursue higher academic goals as measured by Advanced Placement course taking. These positive effects are particularly salient for low-achieving and low-attendance students.

VERSION: April 2019

Engaging Teachers:

Measuring the Impact of Teachers on Student Attendance in Secondary School

Jing Liu and Susanna Loeb

Brown University

December, 2018

Jing Liu (e-mail: jingliu@brown.edu) is a Postdoctoral Research Associate at the Annenberg Institute, Brown University. Susanna Loeb (e-mail: susanna_loeb@brown.edu) is director of the Annenberg Institute and a professor in education and international and public affairs, Brown University. *Jing Liu would like to thank financial support from the National Academy of Education/Spencer Dissertation Fellowship and the Karr Family Fellowship at Stanford's Center for Education Policy Analysis. We are grateful for the substantial help we received from Ben Backes, Lindsay Fox, Dan Goldhaber, Kirabo Jackson, Demetra Kalogrides, Emily Penner, Sean Reardon, Jane Rochmes, Jonah Rockoff, Min Sun, Monica Vasquez, Camille Whitney, and participants in the Family and Education Workshop, AEFP Conference, APPAM Conference, CEPA Seminar, and seminars at Harvard and University of Kentucky. In addition, we extend deep gratitude to several individuals at the school district. We do not mention them by name in order to protect district anonymity.*

Abstract: Teachers' impact on student long-run success is only partially explained by their contributions to students' short-run academic performance. For this study, we explore a second dimension of teacher effectiveness by creating measures of teachers' contributions to student class-attendance. We find systematic variation in teacher effectiveness at reducing unexcused class absences at the middle and high school level. These differences across teachers are as stable as those for student achievement, but teacher effectiveness on attendance only weakly correlates with their effects on achievement. We link these measures of teacher effectiveness to students' long-run outcomes. A high value-added to attendance teacher has a stronger impact on students' likelihood of finishing high school than does a high value-added to achievement teacher. Moreover, high value-added to attendance teachers can motivate students to pursue higher academic goals as measured by Advanced Placement course taking. These positive effects are particularly salient for low-achieving and low-attendance students.

JEL Codes: I21, J44, J45

Key words: value-added models; attendance; non-cognitive outcomes; student long-run success

Introduction

Both anecdotal and systematic evidence points to the importance of teachers for students' long run success. Previous research on effective teachers has focused largely on student test score gains in math and reading in the year in which the teacher teaches the student. This research has shown that a high value-added teacher improves student short-term achievement (e.g.,Clotfelter, Ladd, and Vigdor, 2007; Goldhaber, 2007; Rivkin, Hanushek, and Kain, 2005) and can have long-term impacts on college attendance, income and other adult outcomes (Chetty, Friedman, and Rockoff, 2014). However, a large portion of teacher effects on student long-term outcomes, like college attendance, is not explained by teacher effects on student achievement (Chamberlain, 2013), suggesting that good teachers not only increase students' test scores, but also impact other capacities. As one example, teachers may affect students' school engagement, which can have long-term benefits even if it does not improve test scores in the short run.

An emerging literature sheds light on teachers' impact on students' non-achievement outcomes (e.g., Gershenson, 2016; Jackson, 2018; Kraft, 2017). Most of these studies focus on elementary or middle school level, examining a range of outcomes including psychological traits such as growth mindset and grit (Kraft and Grace, 2015), self-reported self-efficacy and happiness (Blazar and Kraft, 2017), academic motivation (Ruzek et al., 2015), teacher-reported measures of children's social and behavior skills (Jennings and DiPrete, 2010), and full-day absences (Gershenson, 2016; Ladd and Sorensen, 2016). The only study looking at non-achievement measures for high school teachers, Jackson (2018), estimates 9th grade teachers' effects on a composite measure of student GPA, on-time grade completion, suspensions and full-day attendance. Jackson (2018) also shows that non-achievement dimensions of teacher effectiveness can contribute to students' long-run success above and beyond the teachers' impact on student test scores.

The current study extends Jackson (2018) by focusing on teachers' contribution to student class attendance in secondary school, and then linking this new measure to students' long-run outcomes. Research provides evidence that lower attendance results in less learning (Goodman, 2014; Gottfried, 2009, 2010, 2011). Moreover, absence predicts long-term outcomes such as high school dropout, net of other factors including achievement, and is an indicator of student risk such as for drug use and crime (Allensworth and Easton, 2007; Balfanz and Byrnes, 2012; Balfanz, Herzog, and Mac Iver, 2007; Gottfried 2011). While a variety of individual and family factors can lead students to miss school, such as student illness (Romero and Lee, 2007) and residential mobility (Hanushek, Kain, and Rivkin, 2004), factors within the purview of schools such as a positive and safe school environment and an effective, supportive and engaging teacher are also likely to influence absences.

The paper makes four main contributions to this literature. First, we focus on attendance at the secondary school level rather than the elementary level. Since during secondary school students themselves rather than their parents are likely to make the decision of attending classes, attendance in secondary school is more likely to be affected by the student's own perceptions of the teacher than it is in elementary school and is likely, therefore, a more appropriate setting for estimating "teacher effects" on absences. Absenteeism is also higher in secondary school (Whitney and Liu, 2017). On average, secondary students are absent from school an average of three weeks per year (Snyder and Dillow, 2013). Gershenson (2016) focuses on elementary grades and, although Jackson (2018) studies 9th graders, ours is the first study that examines both middle and high school students (7th to 11th graders). Second, the detailed administrative data that we use provide information on whether a student missed each class of each day for either an excused or unexcused reason. Gershenson (2016) and Jackson (2018), as well as other research on student absence, focus on full-day absences as the outcome measure. Because middle and high school students attend classes with multiple teachers, attributing full-day absences to individual teachers is difficult. Class-level absence greatly improves the precision of measuring absenteeism and estimating teacher effects. Moreover, focusing on unexcused rather than the combination of excused and unexcused absences allows us to isolate the types of absences that are more likely a reflection of the student's perceptions of the teacher. A recent study of middle and high school student attendance in one urban school district find that approximately half of all absences from class were due to class skipping on days that students attended rather than to full-day absences (Whitney and Liu, 2017). Third, unlike existing studies that unambiguously treat absences as a continuous variable despite the fact that absences are a discrete count variable (i.e. 0, 1, 2, 3, etc.) and have excessive zeros, we employ a two-level negative binomial model to estimate teacher effects on absences, a method specifically designed for estimation of count variables (Ellison and Swanson, 2016).[1] Finally, and most substantively, using high school graduation, dropout, and Advanced Placement (AP) course taking data, we directly test whether teacher value-added to attendance has predictive power for student long-term outcomes above and beyond teachers' impact on student test scores. Only Jackson (2018) has looked at longer-run outcomes. In this analysis, we also ask whether subgroups of students respond differently to each dimension of teacher effectiveness.

Specifically, this paper answers the following research questions:

[1] We also replicate all the results using an OLS model.

Research Question 1: Variance. To what extent do teachers vary in their contribution to student class attendance?

Research Question 2: Stability. How well does a teacher's value-added to attendance in the current year predict his or her future value-added to attendance, and how does this cross-year relationship compare to that for value-added to achievement?

Research Question 3: Similarity. To what extent are teachers who contribute most to student attendance the same ones who contribute most to student test performance?

Research Question 4: Effects. Does attending classes with a teacher who has high value-added to attendance benefit students in the long run?

The paper proceeds as follows. First, we summarize related literature to motivate the importance of attendance and how teachers can influence student class attendance. We then describe our data and show which student and class characteristics predict absences. In the methods section, we present our identification strategy of estimating teacher effects on student attendance as well as on test scores. Lastly, we describe our approaches for answering the other research questions, present results and robustness checks, and conclude with a discussion of the implications.

Overall, we find that teachers have large effects on student attendance. A student would have approximately 44 percent fewer unexcused absences in math classes and 54 percent fewer in English classes, if she had a teacher who is one standard deviation above the average in value-added to attendance than if she had an average teacher, holding other variables constant. Compared with value-added to achievement, value-added to attendance is similarly stable across years and, yet, value-added to attendance is weakly correlated with value-added to achievement. Compared with having a high value-added to achievement teacher, a high value-added to attendance teacher has stronger effects on a student's opportunity to graduate from high school, and meaningful effects on students' pursuit of higher academic goals such as taking more AP courses. Notably, the effects of value-added to attendance are particular strong for students with lower prior achievement and attendance, while show little impact on students who are at the top of the distribution.

Background

Prior research has documented the critical role of noncognitive skills for a host of long-term socioeconomic outcomes (Heckman, Stixrud, and Urzua, 2006; Cunha, Heckman, and Schennach, 2010). While latent noncognitive skills may be difficult to measure, behaviors, such as attendance and class disruptions, are correlated with well-known psychological measures such as the "Big Five"[2] personality traits and may serve as good proxies for them (Heckman and Kautz, 2013). School attendance is a particularly valuable measure because it can be measured relatively easily and objectively. Psychologists find attendance positively associated with conscientiousness (Duckworth et al., 2007) and negatively associated with neuroticism and low levels of agreeableness (Lounsbury et al., 2004), among other character skills that are valued in the labor market (Heckman, Pinto and Savelyev, 2013).

Quasi-experimental research consistently shows a negative relationship between absences and test scores. Gottfried (2009, 2010, 2011) uses data from the School District of Philadelphia to examine several facets of the relationship between student absences and achievement. Using proximity from students' homes to their school to instrument for attendance and controlling for

[2] The Big Five character skills include Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

school fixed effects, Gottfried (2010) finds a positive relationship between attendance and both grade point average and test scores. Gottfried (2011), alternatively, uses family fixed effects on a longitudinal sample of siblings to control for unobserved heterogeneity in the family environment which might affect both absences and school performance. He finds a negative relationship between absences and achievement even within families. In a more recent study, Goodman (2014) uses snow days as an instrumental variable and finds that each absence induced by moderate snowfall reduces student math achievement by 0.05 standard deviations. Goodman (2014) also finds evidence that absences can cause lower achievement even among non-absent students. The teacher is likely to have a “coordination problem” because when absent students return to school, the teacher may need to allocate instructional time to catching the students up on what they missed (Goodman, 2014).

The prior literature has hypothesized about the role of teachers in encouraging or discouraging absences, though very little empirical work has addressed this relationship directly. Monk and Ibrahim (1984), for example, hypothesize that “if a teacher is weak, or a class is unruly and a student is not learning, the student may respond by being excessively absent.” Ladd and Sorensen (2017), similarly, hypothesize that "effective teachers do more for students on a daily basis than simply imparting a narrow set of reading or math skills…such teachers cultivate character, discipline, and curiosity, and a variety of other capacities.”

Recent research has focused on estimating teacher effects on student social and behavioral outcomes including attendance (Backes and Hansen, 2015; Blazar and Kraft, 2017; Gershenson, 2016; Jackson, 2018; Jennings and DiPrete, 2010; Kraft, 2017; Kraft and Grace, 2015; Ladd and Sorensen, 2016; Ruzek, et al., 2015). Of these, Gershenson (2016) is the only study that focuses specifically on teachers’ impacts on student attendance. The author uses data for 3^{rd} to 5^{th} graders from North Carolina, and an aggregated measure of absences that does not differentiate excused and unexcused reasons. He finds teacher effects on student attendance that are of approximately the same magnitude as effects on achievement. In a similar vein, Blazar and Kraft (2017) estimate teacher effects for 4^{th} and 5^{th} grade teachers on a range of student self-reported measures including self-efficacy in math and happiness. They find that the magnitudes of teacher effects on these measures are similar to that for test scores. These studies find weak or moderate correlations between teacher effects on non-test score outcomes and their effects on test scores, indicating that teaching is likely to be a multi-faceted activity that cannot be captured well by a single outcome measure for students.

The only study looking at non-achievement measures for high school teachers, Jackson (2018), estimated 9^{th} grade teachers' effects on a composite measure of student GPA, on-time grade completion, suspensions and full-day attendance in North Carolina. Jackson (2018) distinguishes itself from other similar work not only by examining high school students but also by directly testing how teacher effects on non-cognitive outcomes affect students' long-run success. He finds that teacher effects on these outcomes have stronger predictive power than their effects on test scores on student educational attainment, such as high school graduation, SAT taking, and intended college going. Effects of teachers on these non-test outcomes are key to explaining their effects on long-run outcomes, particularly for English teachers. The current study extends Jackson (2018) by focusing on class-attendance for the entire secondary grades (except for 12^{th} grade).

Data

Longitudinal administrative data including school years 2003-2004 through 2013-2014 come from a medium-sized urban school district in California. We focus on 7^{th} to 11^{th} graders, excluding 12^{th} graders for two reasons. First, 12^{th} graders do not take standardized tests so we would not be able to estimate teachers' value-added to test performance for that grade. Second, 12^{th} graders are about to graduate and thus have weaker motivation to attend class compared to students in prior grades (i.e., so called "senioritis"), making 12^{th} graders a special population that deserves separate analyses.

The most unusual feature of this dataset is that it has student attendance records for each class on each day and the reasons for absence. During the school years we examine, teachers used a paper scantron to mark a student as absent, tardy or present in each class. For an absent student, a clerk in the school office would mark the student as excused absent if they received a phone call from a parent or guardian providing reasons for absence, otherwise the student was identified as unexcused absent for that class. According to our interview with several administrators in the district, teachers had incentives to report absences accurately. The school did not want them to over report absences because funding was tied to Average Daily Attendance (ADA); however, if a student was recorded as in class when he or she had a discipline issue or was found to be outside of school, that difference could create problems for the school and the teacher. Thus, deliberate misreporting was not perceived as a problem, though a teacher might make mistakes when tracking student attendance. Whitney and Liu (2017) conduct several successful validity checks about the classification of part-day and full-day absence using the same dataset.

Such detailed class-level attendance data are rarely available for researchers. As a result, nearly all the current studies of student attendance use full-day absences with just a few exceptions and none addressing teacher effects.[3] Since part-day absences account for more than half of total absences (Whitney and Liu, 2017), ignoring part-day absences may result in significant error for estimating days when students do not attend particular classes and may bias estimates of value-added to attendance as well as result in less reliable measures, especially when part-day absences are non-randomly distributed among students with different characteristics. Class-level attendance data also allow us to directly estimate teacher effects on the specific student absences they are responsible for. Since we have information on whether absences are excused or unexcused, we are able to focus on unexcused absences which are more discretionary for students and thus more likely to be affected by teachers.

We combine several databases to construct our final sample. First, we match classes in the attendance dataset to their corresponding subject area. We focus our analysis on five core subjects – math, English language arts (ELA), science, social studies, and foreign languages – because non-core subjects like physical education have relatively fewer teachers. Second, we link student attendance data to a rich set of student demographic variables including race/ethnicity, gender, English learner status (EL), special education status, and gifted status. Third, we add in student test scores from California Standards Tests (CST). Students in grades two to eleven were required to take these state mandated tests during the years of our study. Although we also have test scores for science and history, we only use math and ELA test scores in this study because science and history were not tested in each grade. We link teachers to students using a combination of class identifier, school, grade, period, and teacher ID, which in turn allows us to construct classroom level covariates. For our last research question, we merge

[3] One prior study of absences uses the average number of absences to each class but does not examine effects of teachers on this measure (Cortes, Bricker, and Rohlfs, 2012).

in student high school graduation status, whether they dropped out before 12th grade[4] and their AP course taking behavior.

In secondary school, a student can have multiple teachers and classes in a subject but only one test score in that specific subject area, which creates a problem for how to construct samples for estimating different value-added scores. For the main analysis, we estimate value-added to attendance and achievement for math and English teachers on the same sample that we fully describe below, though we run specification checks using a range of other samples. We choose this restricted sample so that we can compare value-added measures on attendance and achievement without the concern that these two measures are estimated using different samples. To create this sample, we constrain the data in several ways. We begin with observations at the student-class period-semester level. We only use students who have one teacher in a subject for the entire school year. This restriction cuts nearly a quarter (22.6%) of our sample[5] but removes the difficulty of disentangling teacher effects on student test scores when multiple teachers are present. This restriction reduces the generalizability of our sample somewhat, though fits with our purpose of understanding value-added to attendance. We also drop student-class period-semester observations if one is absent from more than 50 percent of their total classes because

---

[4] The district gave us those long-term outcomes up to school year 2014-2015. For later cohorts (e.g., those who were 7th graders in 2012-2013) that we do not have data to observe their graduation and dropout, we assign missing values to these outcomes. For all 7th graders we can observe in our sample, 56.51% graduated from high school, and 28.39% dropped out before 12th grade. These numbers are 59.30% and 22.50% for 8th graders, 68.40% and 16.51% for 9th graders, 81.39% and 8.27% for 10th graders, and 91.18% and 3.37% for 11th graders. For those who neither graduated nor dropped out, some are transitional school students (8th to 9th grade) who did not return to the district and did not submit a school enrollment application for the subsequent year.

[5] 24.21 and 15.54 percent of students have more than one teacher in English and math, respectively.

students with extremely high absence rates are likely to be absent due to reasons beyond a teacher's control. We thus also drop observations when the student has less than ten valid attendance marks in a class per semester.[6] We also exclude classes with fewer than five students because we would lack precision when estimating teacher effects for such small classes. For the comparison of value-added to achievement and value-added to attendance, we limit the sample to teachers for which we can compute both measures. These restrictions drop an additional 12 percent of the sample.

Table 1 reports the descriptive statistics at the student, classroom, and school level. The first four columns report characteristics of students in all five subjects for both the full sample and the restricted sample,[7] while the next set are math specific and the last set are for ELA. Overall, we lose about 20 percent of student-year observations using our analytical sample, resulting in 185,000 student-by-year observations and 8,900 teacher-by-year observations. Compared with the full sample, the analytical sample has slightly less black and Hispanic students. The average test scores are lower and number of absences are also higher, suggesting that we are using a slightly more advantaged group of students, though overall, the samples are quite similar.

Our analytical sample includes slightly more male students than females. One salient feature is that students are racially diverse. About 50 percent of the students are Asian, slightly over 20 percent are Hispanic, and about 10 percent are black. The racial composition is similar

---

[6] Invalid attendance marks refer to those classes that are inactive, have no record of attendance, or have attendance marks that are miscoded.

[7] We constrain this sample using the same criteria as what we do for the sample used to estimate both value-added to attendance and value-added to achievement, i.e., every student only has one teacher in a specific subject-year, each class has more than five students, and each student has less than 50% of total class absences.

for math and ELA classes compared with the whole analytical sample, with slightly more Asian students in math classes. Given this racial composition, it is not surprising that the percentage of EL is about 20 percent. Since we eliminate students with more than 50 percent of unexcused absences as well as those with multiple math or ELA classes in the same semester, who tend to be academically weaker than other students, the average (standardized) test scores on both math and ELA in our sample are slightly higher than zero. The classroom level and school level statistics are similar to those at the individual level.

On average, students have 3.04 unexcused absences for a math class and 2.96 unexcused absences for an ELA class per school year,[8] accounting for 3.96 percent and 3.84 percent of the total class meetings, respectively. For both subjects, the average excused absences are about half of unexcused absences for a class. The standard deviations are much bigger than the means for both excused and unexcused absences, indicating highly skewed distributions for both variables.[9]

[Table 1 here]

Both student and class characteristics can influence students' decision to attend a class. To better inform our value-added estimation, we use regression analysis to examine how these factors predict unexcused absences.[10] Table 2 synthesizes the results. The dependent variable here is the rate of unexcused absences for a class. In the first two columns, we report results using data for all subjects. In columns three to six, we conduct separate analyses for math and

[8] Students might take multiple classes in a subject in a school year. We report the averages across all class-periods in the corresponding subject in a school year.

[9] We calculate total class meetings for each student-class period cell by aggregating all the unexcused, excused, tardy, and present attendance marks. Classes on average have about 76 meetings in a semester, a 15-week span assuming students met every day. While the school year is 180 days, some classes do not meet every day, particularly at schools with non-traditional schedules. In addition, on some days students in a class may not meet due to special activities such as school-wide assemblies.

[10] For a more comprehensive examination, see Whitney and Liu (2017).

ELA. The first model contains only the reported variables, while the second includes school-by-year fixed effects, so that the comparison of students and classes are within schools in a given year.

Across different subjects and model specifications, female students have significantly fewer unexcused absences than males, but the size of the differences are quite small at about .002 to .003. Differences between ethnic groups, in contrast, are quite substantial. Compared with Asian students (the group left out of the model for comparison), black students have an average unexcused absences rate 6.3 percentage points higher, according to the most conservative estimate. Hispanic students have substantially lower unexcused absence rates than black students but exceed the rates for white students and students from other ethnic groups, each of whom have higher average rates than Asian students. Unexcused absences also differ by grade level. Higher grade levels generally have more unexcused absences, with a large jump between grades 8 and 9 and then relative stability between grades 9 and 11.

Class characteristics also predict student attendance. One important factor is the timing of the class. Most schools have seven total periods, while a few have a zero or eighth period. We group those periods as a separate group. As expected, students skip the first class in a day more than later ones and are second-most absent from their seventh-period class. Class subject is less important to the number of unexcused absences. Our results show that ELA classes have significantly fewer unexcused absences than math, science, social studies and foreign language classes, but the differences are small.

[Table 2 here]

To facilitate constructing our value-added measures, we aggregate our data from student-class period-semester level to student-teacher-year level, which allows us to estimate teacher effects on both test scores and attendance using the same dataset. Although a student has only one test score in a subject in a year, students can take more than one class-period with a teacher in a subject, so we aggregate absences for each student-teacher-subject-year combination.[11] This method allows students to have different *exposure times* or total class meetings with a teacher, and, thus, the total number of absences after aggregation are not directly comparable between students anymore. In what follows, we provide a detailed explanation on how our model addresses this issue. For class-level controls, we calculate the average classroom characteristics for all class-periods a student took with a teacher in a certain subject in a year.

Empirical Strategy

Estimating Value-Added to Attendance: We estimate a two-level negative binomial regression model (NBRM) to construct value-added measures for teacher's impact on student attendance in the process and estimate the variance in this measure across teachers. Prior studies that use student test scores as outcome variables generally employ an Ordinary Linear Square (OLS) model with teacher or teacher-by-year fixed effects to estimate value-added. The NBRM is better suited than OLS to model teacher effects on attendance given that attendance is a count variable and has excessive zeros, though in some cases NBRM and OLS will provide similar results even with count data. We test the relative merits of the approaches for our data by replicating the analyses using OLS models and find that OLS performs worse than NBRM, which we explain in detail later. As Graph 1.1 shows, the distribution of unexcused absences is extremely skewed. Around 40% of the values are zeros at the student-class period level for math

[11] If counting one class period-semester as a class (so Algebra 1 in fall and algebra 2 in spring are counted as two classes), 33.22% students have just one class with a teacher in a subject in a year. 63.50% have two classes.

classes. A similar pattern holds if we plot the percentage of unexcused absences over total class meetings for all five subjects at the student-school year level; collapse the data to class period level; or collapse the data to teacher-year level.

[Graph 1.1-1.4 here]

The NBRM belongs to the family of models which deal with counts as dependent variables. The NBRM is an extension of the Poisson regression model, adding one more parameter to account for over-dispersion in the dependent variable, which allows the variance to exceed the mean.[12] The NBRM allows the number of events to have different exposure times and thus can account for students who have the same teacher for different meeting times in a year. We embed the NBRM into a two-level random intercept framework to estimate teacher effects. A two-level random intercept model estimates the variance of value-added directly and provides empirical Bayes estimates of individual teacher effects (McCaffrey et al., 2004).[13]

The greatest challenge to estimating teacher effects is that students may not randomly sort to teachers. Several studies show that controlling for student prior test scores eliminates most of the sorting bias when creating measures of value-added to test performance (Chetty, Friedman, and Rockoff, 2014; Kane and Staiger, 2008). To reduce the possibility of bias from

---

[12] We run a simple test to show that the NBRM outperforms the PRM in our setting. We regress student unexcused absences on basic student, class, and school covariates with both models. Then we predict the expected number of unexcused absences given the results of these two models. If there is a smaller difference between the observed value and predicted value for the NBRM compared with the PRM, it suggests the NBRM fits the data better.

[13] A common debate in the value-added literature is whether teacher effects should be treated as fixed or random. In our case, although Hausman, Hall, & Griliches (1984) propose a conditional likelihood method for negative binomial regression with fixed effects, Allison & Waterman (2002) show that it does not qualify as a true fixed effects because time-invariant covariates are allowed in their model and can result in a non-zero coefficient on those covariates. This problem arises because the model allows for individual-specific variation in the dispersion parameter instead of in the conditional mean (Rabe-Hesketh and Skrondal, 2008). We thus choose to embed the NBRM into a two-level random intercept framework to estimate teacher effects.

within-school sorting, we control for student prior absence rates in the same subject as well as in other subjects, in addition to controlling for student prior test scores. Unlike when calculating teacher effects using data on elementary and middle school students, simply controlling for prior test scores and absences may not fully eliminate selection bias at the high school level because of academic tracks and unobserved track-level treatments (Jackson, 2014). The school district we study does not use formal academic tracking. However, students in secondary school chose which math courses to take and, as a result, take different math tests at the end of the year. We use interactions of grade and the test the students took in that grade as well as in the prior year to control for possible selection of teachers and students into different courses. Following Jackson (2018), we formally test selection on both observables and non-observables to further assess whether our controls sufficiently remove bias due to sorting. We do not find evidence of substantive bias for value-added to attendance, which we discuss in the Robustness Checks section.

We pool all grades together and estimate the following models separately by subject:

Level 1:

$$E(Abs_{ijt}) = \mu_{ijt} = \exp\left(X'_{ijt}\beta + \theta_{jt} + \varepsilon_{ijt} + \ln\left(ET_{ijt}\right)\right)$$

where $$Abs_{ijt}|\mu_{ijt} \sim Poisson(\mu_{ijt})$$

and $$\exp\left(\varepsilon_{ijt}\right)|\theta_{jt} \sim Gamma(r_{ijt}, p_{ijt})$$

and $$Cov(X_{ijt}, \theta_{jt}) = 0$$

and $$Cov(X_{ijt}, \varepsilon_{ijt}) = 0.$$

r_{ijt} and p_{ijt} are two parameterizations of conditional overdispersion. Specifically,

$$r_{ijt} = 1/\alpha \text{ and } p_{ijt} = \frac{1}{1+\alpha * \exp\left(X'_{ijt}\beta + \theta_{jt}\right)}$$

Level 2:

$$\theta_{jt} = \gamma_{00} + u_{jt}$$

where $$u_{jt} \sim N(0, \psi)$$

In this model, the variation driving the estimation comes from teachers who taught students that share similar achievement, attendance, and demographic characteristics and were in the same types of courses, adjusting for differences across grade levels and years. In level one, $E(Abs_{ijt})$ or μ_{ijt} indicate student i's expected unexcused absences with teacher j in school year t. X_{ijt} represents a variety of student, classroom, and school characteristics. Appendix A gives a full list of controls. θ_{jt} is a random effect for teacher j in year t, which is the teacher-by-year value-added estimate of interest. Alternatively, we estimate value added on the teacher level by replacing θ_{jt} with θ_j, which directly provides variance estimates across teachers. ε_{ijt} is a random error that results in over-dispersion, the reason for choosing the NBRM model. ET_{ijt} indicates the exposure time (i.e. total class meetings), for student i with teacher j in school year t. By adding this exposure variable, we control for differences in exposure times, with the coefficient constrained to one (Long and Freese, 2014). In level 2, our teacher-by-year effect θ_{jt} follows a normal distribution with mean equal to 0.[14]

In our preferred model, we do not include school fixed effects, although we also run a version of the model with school fixed effects to compare their performance. First, the inclusion of school fixed effects reduces stability and creates additional noise in value-added estimates because only teachers who move between schools identify the school effects (Mihaly et al.,

[14] The variance of u_{jt} (i.e., ψ) is a $q \times q$ variance matrix $\sum$. The conditional distribution of $\boldsymbol{Abs_{jt}}=\{Abs_{1jt}, \cdots, Abs_{njt}\}'$ given random effects u_{jt} and the conditional overdispersion parameter α is $f(\boldsymbol{Abs_{jt}}|u_{jt}, \alpha) = \prod \frac{\Gamma(Abs_{ijt}+r_{ijt})}{\Gamma(Abs_{ijt}+1)\Gamma(r_{ijt})} p_{ijt}^{r_{ijt}}(1-p_{ijt})^{Abs_{ijt}}$. The likelihood function for teacher j in year t is $\mathcal{L}_{jt}(\beta, \sum, \alpha) = (2\pi)^{\frac{1q}{2}}|\sum|^{-\frac{1}{2}} \int f(Abs_{jt}|u_{jt}, \alpha) \exp\left(-u_{jt}' \sum^{-1} u_{jt}/2\right) du_{jt}$.

2013; Mansfield, 2015). In our case, only about 8.5% teacher switched between schools within the district during the period we examine, which creates much noise in value-added scores. Second, the literature on teacher effects on achievement find that once controlling for prior performance, adding school fixed effects does little to further remove bias (Koedel, Mihaly, & Rockoff, 2015). Nevertheless, school-level policy may influence student attendance and teachers' ability to reduce student absences. Thus, while we estimate the value-added model without school fixed effects, we include school fixed effects in models estimating how value-added to attendance affects student short- and long-run outcomes. Thus, these estimates use only within-school variation.

Given the nonlinear nature of the model, we can interpret teacher effects as the percentage change of the expected number of unexcused absences. By computing the equation below, we get the percentage change in the expected number of unexcused absences for a student with a teacher who has a value-added of one standard deviation above the average, compared with the result assuming the student has an average teacher, holding other variables constant.

$$\frac{E(Abs_{ijt}|X_{ijt},\psi^{1/2})}{E(Abs_{ijt}|X_{ijt},0)} = \exp\,(\psi^{1/2})$$

The result is given by 100 x $(\exp\left(\psi^{1/2}\right) - 1)$.

We predict the Empirical Bayes (EB) estimates of the teacher-year effects by using the means of the empirical posterior distribution with the estimated model parameters including $\hat{\beta}$, $\hat{\alpha}$, and the variance components of ψ.[15] We then standardize these EB estimates to have a mean of zero and a standard deviation of 1. Since the dependent variable is unexcused absences, a bigger value in these EB estimates indicates a bigger effect on increasing unexcused absences.

---

[15] For a complete introduction of this procedure, see Skrondal and Rabe-Hesketh (2004, chap. 7)

To ease interpretation, we convert them to *value-added to attendance* by multiplying all the EB estimates by -1.

We test the robustness of our results using excused absence as an outcome variable to conduct a form of a placebo test. In theory, students need legitimate reasons, such as sickness, to have an excused absence, which should be free from a teacher's influence. In practice, excused absences may be fungible, for example, if students are more likely to feign sickness to schedule doctors' appointments or other appointments during classes in which they are not engaged. Nonetheless, unexcused absences are likely to be more affected by teachers than are excused absences, and, thus, the estimated variance of value-added to unexcused absences should be larger than that for excused absences.

Estimating Value-Added to Achievement: We estimate value-added to achievement so that we can examine how the two measures of teacher effectiveness correlate and compare their stability. We adopt a widely used strategy which estimates teacher or teacher-by-year fixed effects, accounting for student math and reading test scores in the prior year. An experimental study shows that this model outperforms other models of teachers' value added to student test scores (Guarino et al., 2015). We compute empirical Bayes estimates after running the fixed effects model by summarizing the estimated standard errors to estimate the sampling error variance and then shrinking the estimates by a signal-to-noise ratio based on this sampling error variance.

Results

RQ1: Variance. By running NBRM separately for each subject, we obtain a standard deviation of value-added to attendance for each of math, ELA, science, social studies, and foreign languages.[16] Table 3 gives the results, reporting both teacher level and teacher-by-year level estimates since both come into play in subsequent analyses. The first column shows the raw standard deviations. Predictably, teacher level estimates have smaller variances than teacher-by-year estimates for each subject.

Because we use the number of unexcused absences as dependent variables, these standard deviations do not provide an intuitive interpretation of the magnitude, nor can we compare them directly to value-added to achievement. Instead, in column 2 of Table 3, we report the incidence rate ratio (IRR) of one standard deviation of value-added to attendance. The magnitude here is easily interpretable. For example, a student would have 44.3 percent fewer unexcused absences in math classes if she had a teacher who was one standard deviation above the average than she would if she had an average teacher, holding other variables constant. The magnitude equals 0.27 SD of or 1.35 unexcused absences. This number is greater, 54.1 percent, for English classes (0.33 SD; 1.60 unexcused absences).[17]

[Table 3 here]

As expected, we find that the magnitude of the variance of value-added to attendance when using excused absence is smaller than for unexcused absences. Specifically, the standard deviation is 0.23 for math (compared with 0.37 for unexcused absence), and 0.24 for ELA

---

[16] In Appendix B, we report the regression results for estimating math teachers' value-added to attendance. In appendix C, we report both the variances and the stability of value-added estimates after adding school fixed effects to our base specification. As expected, value-added to attendance has a much smaller variance and is less stable after adding school fixed effects.

[17] We also estimate a model controlling for school fixed effects, so we are only comparing teachers within schools, though, as discussed above the approach has drawbacks. The standard deviation is 0.282 (IRR=1.326) for math teachers and 0.348 (IRR=1.417) for English teachers, both slightly smaller than the results without school fixed effects. Because school fixed effects models identify school effects through the relatively small sample of movers, we prefer a model without school fixed effects, and focus on this preferred model in this paper.

(compared with 0.43 for unexcused absence), suggesting that unexcused absence is more malleable and teachers have a greater impact on it.

RQ2: Stability. To investigate the stability of value-added to attendance, we conduct two analyses. First, we generate transition matrices to examine how teachers' quintile rankings change from the first two years we observe them to their third through fifth years. Specifically, we compute teachers' quintile ranking by taking the average of each teacher's first two years' value added and also the following three years. If a large proportion of teachers stay where they are initially in their third to fifth years or move very little, we have evidence to say that value-added to attendance is a relatively stable measure. Although the transition matrices provide us an intuitive way to measure the stability of value-added to attendance, it does not offer a succinct measure of how well early value-added predicts future value-added. Thus, we conduct a second analysis which regress value added of a future year (3, 4, or 5) on their first two years of value added. The adjusted R-squared statistics measures how much variation is explained by teachers' early years' effectiveness (Atteberry, Loeb, and Wyckoff, 2015). To benchmark the results, we do the same analysis on value-added to achievement so that we can compare the measures. Throughout these two analyses, we limit our analytical sample to teachers who have at least five years of value-added on attendance and on achievement.

We find substantial stability in value-added measures for teachers over time. Table 4.1 reports the quintile transition matrices for value-added to class attendance.[18] About 67 percent of teachers who are in the lowest quintiles in terms of their average value-added to attendance during the first two years we observe them (the least effective ones), stay in the bottom two

---

[18] We compute teachers' ranking quintiles by subject, but in those transition matrices we combine math and English teachers into one table.

quintile in the following three years; 78 percent of the initially top teachers stay in the top two quintiles.

Table 4.2 for comparison gives the corresponding transition matrices for value-added to achievement.[19] Value-added to achievement is approximately as stable as value-added to attendance, with 70 percent of the lowest quartile teachers remaining in the lowest two quintiles and 79 percent of the highest quartile teachers remaining in the highest two quintiles (compared with 67 percent and 78 percent for attendance).

[Table 4 here]

Transition matrices have drawbacks as measures of stability because they do not capture variation within the quintiles. To further measure stability, we use regression analysis to measure how early years' effectiveness predicts future years' performance. Table 5 reports the adjusted R-squared from different specifications. The first row of the table reports the adjusted R-squared when we regress value-added to attendance in year 3, 4, 5 and the average of all three years on the first year and/or second year of available value-added measures. The upper half of the table shows results using value-added to attendance, and the lower half, using value-added to achievement.

In keeping with the transition matrixes, the regression analyses show substantial predictive power for the value-added to attendance measures. This predictive power is stronger than for value-added to achievement for math teachers, but not for English teachers. The first two years of value-added to math attendance explains 39.7 percent of the variance in the average value-added in years three through five. This figure for attendance is 22.6 percent for English

[19] After adjusting for measurement error, the true standard deviation of value-added to achievement is 0.17 for math, and 0.10 for English.

teachers. In comparison, the percent explained for achievement is 36.3 percent for math and 28.0 percent for English.

[Table 5 here]

RQ3: Similarity. Our third research question asks how correlated measures of value-added to attendance are to measures of value-added to achievement. We use both Pearson correlation and Spearman rank correlation to examine this question. We disattenuate the Pearson correlations by dividing the correlations by the square root of the product of the reliabilities of each value-added measure in each subject. We expect to see a stronger correlation when using teacher value added than using teacher-by-year value added because the teacher level estimates use information from multiple cohorts of students and will be less prone to measurement error.

Alternatively, we run a joint multi-level model to estimate our two value-added measures simultaneously so that we can estimate the covariance directly. Previous research has used a similar approach to examine whether the same teacher has differential effects when teaching different subjects (Fox, 2016) or different types of students (Loeb, Soland, and Fox, 2014). Although this model does not allow us to use a NBRM framework anymore, it has the benefits of reducing sampling errors. Appendix D gives a more detailed description of this model.

Another approach to this question is to regress student outcomes (i.e., test score and rate of unexcused absence) on value-added to achievement and value-added to attendance.[20] If these two measures capture distinct dimensions of teacher ability, we would expect to see no impact of value-added to achievement on attendance, and value-added to attendance on test scores. To avoid "mechanical endogeneity" of our value-added measures as discussed by Chetty et al.

---

[20] To ease interpretation, we run linear regressions for both outcomes, though we use a non-linear model to estimate value-added to attendance.

(2014) and Jackson (2018), we estimate "leave-year-out" value added by using all data but not the year when the focal student has the teacher (i.e., Jackknife estimates). We standardize those value-added estimates using the "true" standard deviations of teacher effects estimated in RQ 1 and RQ 2.[21]

Here we report correlations between value-added to attendance and value-added to achievement. We have both teacher and teacher-by-year measures but we prefer the teacher level since it is less vulnerable to measurement error. While overall the correlations are small, it is a bit stronger for math than for ELA. After adjusting for reliabilities,[22] the Pearson correlation is 0.115 for math, and 0.082 for ELA. Correspondingly, the Spearman rank correlation for math teachers is 0.132 and for English teachers 0.012.[23] We further test these results by using a joint model to directly estimate the covariance of two kinds of value-added scores. The resulting correlation is 0.063 for math and 0.070 for ELA. As a reference, Gershenson (2016) reports near zero correlations (Spearman rank correlation is 0.04 for math teachers and 0.02 for English teachers) for elementary teachers. Pooling together both math and ELA, Jackson (2018) reports a Pearson correlation of 0.097 for 9th grade teachers.[24] Our results are consistent with the literature

---

[21] For the two-level negative binomial model, the variance of teacher value-added to attendance is directly estimated. For the fixed effects model used to estimate value-added to achievement, the true variance equals the observed variance minuses the variance of errors.

[22] The reliability of value-added to attendance is 0.82 for math, and 0.79 for ELA. The reliability of value-added to achievement is 0.89 for math, and 0.70 for ELA.

[23] If we use value-added measures from a model with school fixed effects, the correlations are similar to the results here. Specifically, after adjusting for measurement error, the Pearson correlation is 0.135 for math, and 0.015 for ELA. Correspondingly, the Spearman rank correlation for math teachers is 0.174 and for English teachers is 0.015.

[24] Both Gershenson (2016) and Jackson (2018) originally report negative cross-domain correlations because they do not convert teacher effects on absence to teacher effects on attendance. We change the direction here to ease comparison.

in terms of showing teacher's effectiveness as multi-dimensional, as suggested by the low correlations across measures.

[Graph 2 here]

To further address this question, we directly regress student current outcomes on both the value-added to achievement and value-added to attendance of their teachers. We use out-of-sample estimates of value-added to avoid "mechanical endogeneity," standardizing them by the "true" standard deviations of teacher effects estimated using all years of data. The standard deviation of value-added to achievement is 0.17 for math, and 0.10 for ELA. The standard deviation of value-added to attendance is 0.37 for math, and 0.44 for ELA. In these models, we include school fixed effects in order to eliminate the time-invariant factors within schools that could affect both measures of teacher effectiveness and student outcomes.

Table 6 presents the results. As columns (1) and (5) show, the leave-year-out estimates of teacher effects for one outcome have a strong impact on that outcome. A one standard deviation increase in value-added to achievement improves student test scores by 0.08 standard deviation (p-value<0.01), and a one standard deviation increase of value-added to attendance reduces a student's unexcused absence rate by 0.79 percentage points (p-value<0.01). Columns (2) and (4) indicate that a more effective teacher in increasing student test scores can also reduce student absences, and vice versa, though the magnitude is much smaller compared with those from columns (1) and (5). This result is expected given the weak positive correlation between the two measures of teacher effects. When including both value-added estimates in the same regression, conditional on value-added to achievement, value-added to attendance does not demonstrate an impact on test scores, and both the magnitude and significance of value-added to achievement stay approximately the same. Similar results hold when using unexcused absence rates as the outcome. These results further confirm the weak correlation of our two measures of teacher effects, which measure largely distinct dimensions of teacher ability.

[Table 6 here]

RQ4: Effects. —To examine the effects of high value-added to attendance teachers on longer-run outcomes, we use as outcome measures high school graduation, dropping out from high school before 12th grade, and the total number and credits earned for AP courses in 12th grade.[25] The two sets of measures – one focused on completion and one on higher-level course taking – allow us to examine outcomes for students at different parts of the academic distribution. The completion margin is more salient for students who are marginally engaged with schools and, on average, are likely to have lower achievement and attendance, while AP taking is more salient for highly engaged students who are choosing between more and less challenging coursework but, on average, have higher achievement and attendance.

To construct this dataset, we pool math and ELA classes for all 7th to 11th graders. Under this data structure, each student has one outcome but multiple observations (because of multiple subjects and grades). We account for the correlation of observations by clustering the standard errors at both student and teacher levels. We regress our dependent variables on the standardized leave-year-out value-added to achievement and value-added to attendance separately and then together. Of particular interest is whether adding value-added to attendance affects student outcomes in the long run, after controlling for value-added to achievement. In all the models, we control for baseline covariates as what we did in RQ 1, including student demographics, lagged

---

[25] In our sample, 53.36% of AP courses are taken in 12th grade, and 37.97% are taken in 11th grade. We only use AP courses taken in 12th grade to avoid mechanical endogneity since we are using 7th to 11th grade teachers.

test scores and attendance, lagged and current academic “tracks” (test types), classroom and school characteristics. In addition, we include school fixed effects to account for time-invariant school characteristics that could independently affect value-added and student outcomes.

These models allow for the exploration of heterogeneity. First, to explore potential nonlinearity, we add squared terms for both value-added to attendance and value-added to achievement. These analyses provide insights into whether the effects of low value-added teachers on longer-term outcomes are stronger than the effects of high value-added teachers and whether these patterns differ for the two dimensions of teacher effectiveness. Additionally, we classify students into thirds based on their prior achievement and attendance and examine the heterogeneity of value-added measures on students with different performance levels. These analyses provide insights, for example, into whether students who are on the margin of dropping out benefit more from teachers who can keep them at school, while students who are already academically advanced need teachers with a different skillset to motivate them to pursue higher learning goals, such as AP courses.

Table 7 presents the first set of results, pooling data for math and ELA for 7th to 11th graders. The upper panel of table 7 reports results using high school graduation and dropout as the outcome variables. Teachers with high value-added to attendance increase students’ probability of graduating from high school and reduce their chance of dropping out before grade 12, independent of their effectiveness in increasing student test scores. Although value-added to achievement has a positive coefficient for high school graduation, it is insignificant. In contrast, value-added to attendance shows significant impact on high school graduation. Specifically, a one standard deviation increase of value-added to attendance improves a student’s probability of high school graduation by 0.7 percentage points. When putting both measures in the same regression, the coefficients maintain similar magnitude and significance. The results for value-added to attendance are similar when using dropout as the outcome variable. A one standard deviation increase in value-added to attendance reduces a student's probability of dropping out before 12th grade by 0.3 percentage points, with no discernable effect of test-score value added with or without value-added to attendance in the regression.

The story is somewhat different when using AP course taking as outcomes. Value-added to achievement and value-added to attendance both have significant and positive impact on the number and earned credits of AP courses. Here, however, the effects are smaller for value-added to attendance. Specifically, a one standard deviation increase in value-added to achievement increases the number of AP course taking by 0.02 and AP credits by 0.10. These numbers are 0.01 and 0.06, respectively, for value-added to attendance.

[Table 7 here]

To investigate potential nonlinearity of teacher effectiveness, we examine long-term outcomes by adding squared terms for both measures. Table 8 provides the results. In column (1), we find a negative and significant coefficient for the squared term of value-added to attendance, providing evidence that the impact of value-added to attendance on graduation diminishes as it becomes larger. That is, the effects of teachers' value-added to attendance on long-run outcomes are driven by the negative effects of particularly ineffective teachers. These non-linearities for value-added to attendance are not as clear for the other outcomes, with point estimates in the same direction but of smaller magnitude and statistically insignificant. In contrast, the squared terms of value-added to test scores are positive and significant when using both the number and earned credits of AP courses as the outcome variables, indicating that students benefit even more from a high value-added teacher than they are harmed by a low value-added teacher. Here the effects of teachers' value-added to achievement on long-run outcomes are driven by the positive effects of particularly strong teachers. Given the variation in these relationships across the outcomes, the results are suggestive but not definitive.

[Table 8 here]

To examine the potential heterogeneous effects across students, we categorize students by their prior year's performance. Specifically, we run separate regressions for students who are in the bottom and top thirds of attendance and math scores the year before they have a specific teacher. Table 9 reports the results.

Teachers' value-added to attendance teachers is particularly important for low attenders and low achievers. For students who are in the bottom third of attendance or math scores, value-added to attendance has a positive and significant coefficient for all the long-term outcomes, suggesting students who are on the completion margin benefit substantially from high value-added to attendance teachers. In contrast, while high value-added to achievement teachers significantly contribute to AP-related outcomes for low-attendance and low-achievement students, they do not affect the completion outcomes for these students, with the exception of graduation for low-achieving students and here the point estimate is substantially smaller than for value-added to attendance. These results provide some evidence that teachers who improve test scores help less-engaged students academically, but they do not keep students in school. On the other hand, teachers who improve attendance of these students impact both their academics and their school retention.

The results differ for students who are in the top third of attendance and math scores. Since these students, especially high-attendance students, are not missing a lot of school and, likely, have a low risk of dropping out, it is not surprising that we observe little relationship between either type of value added and graduation or dropout of high-attendance students. Value-added to attendance predicts school completion at marginal significance for high-achieving students, which is some indication that high-achieving and high-attending students are not synonymous. In contrast, value-added to achievement positively and significantly predicts AP courses, with estimates nearly three to four times as large as in models using low attenders and low achievers, indicating that top students can benefit substantially from teachers who can help them learn. Value-added to attendance does nothing to support AP course taking for high-attending or high-achieving students. Engagement may not be a factor for these students.

[Table 9 here]

Taken as a whole, our results confirm the multi-dimensional nature of teacher effectiveness. A teacher who has high value-added to attendance may be able to engage students in class and motivate the student to pursue higher academic goals, but this impact is much more salient for students with low-attendance and low-achievement. In contrast, teachers with high value-added to achievement can help students to pursue higher academic goals, but do little to improve graduation and reduce dropout for students prone to missing class.

Robustness checks. The two-level NBRM accounts for the count nature of attendance and, as a result, can work with excessive zeros in the outcome variable. Nonetheless, OLS is a standard and robust to substantial non-normality in the dependent variable. Thus, we rerun all analyses using similar value-added measures based on OLS regression and mirroring our value-added to achievement measures. The results of these analyses indicate that NBRM indeed outperforms OLS when we link value-added to attendance to student outcomes. While many of the results from NBRM hold up using OLS, some do not. Intuitively, an OLS prediction always generates a negative residual for zero values while NBRM does not have this issue. As a result, some teachers with many students with zero absences who receive positive value-added scores in NBRM end up with negative scores in OLS. After eliminating these teachers who are sensitive to the modeling procedure, the results using OLS and NBRM are consistent (see Appendix E for the details).

In keeping with value-added measures in prior studies, our value-added measures for both attendance and achievement are based on models that adjust for selection of students into teacher using controls for observables. Given this approach, bias from selection on both observables and unobservables is possible. The assumption of our analyses is that conditional on the controls in our specification, students are not systematically sorted to teachers. To test the validity of this assumption, following Chetty et al. (2014) and Jackson (2018), we first use twice lagged student characteristics to predict all the long-term outcomes.[26] Using predicted outcomes and conditional on all student, class, and school characteristics excluding those used in the prediction, we should not observe any significant association between the estimated teacher value-added (leave-year-out estimates) and the predicted outcomes. As shown in Appendix Table F1, the significances of value-added to attendance disappear for predicted graduation and dropout, providing some evidence that the main model adjusted successfully for the observables. Although the coefficients are significant for predicted number of AP courses and earned AP credits, the magnitudes are so small that they suggest very little selection in our model. For value-added to achievement, we observe significant coefficients for predicted graduation and dropout, but the directions suggest underestimate, instead of overestimate, of effects. Similar to

---

[26] This approach effectively limits our sample to students who have twice-lagged controls and only 8th to 11th graders. The student characteristics used here include test scores and absence rates for both math and ELA classes, days of suspension, race, gender, special education status, gifted status, and English learner status.

the coefficients on value-added to attendance, the coefficients on value-added to achievement for predicted AP courses are very small and only marginally significant. Overall the results provide evidence that our strategy largely eliminates selection on observables.

We cannot directly test whether our estimates are biased due to selection on unobservables. However, following Jackson (2018), we can assess selection on unobservables by comparing estimates based on two distinct sources of variation. The first strategy relies on school-by-cohort fixed effects. Since Jackson (2018) only uses 9th graders, he uses school-by-year fixed effects. Here we modify his approach by using school-by-cohort fixed effects. This approach should be robust to any school-level policies and shocks that affect all students in a school cohort, since our estimating variation comes from within school-cohort. The second strategy uses a Two-Stage Least Square estimator, using variation in average estimated teacher value-added scores across cohorts within a school. This Instrumental Variable approach is robust to student selection to teachers within a school but is susceptible to selection across schools. If these two distinct identification strategies provide similar results, then we have some additional evidence that our estimation strategy is not biased due to unobservables. As shown in Appendix Table G1, the overall magnitude and significance is remarkably consistent with Table 7, especially for value-added to attendance. We find no evidence of bias due to selection on unobservables.

Discussion and Conclusion

Students in secondary school skip many classes even when they are in the school. Approximately half of the days that they are not in a specific class, they attend other classes (Whitney and Liu, 2017). In this paper, we create measures of middle and high school teachers' individual contribution to student engagement as measured by student class-by-class attendance, asking to what extent teachers vary in their ability to get students to come to class and how much this variation also leads to differential long-run outcomes for students. An extension of Jackson (2018), our study is only the second study that is able to estimate teachers' effect on student non-test score outcomes and then link this measure to student long-run outcomes. We find substantial variation across teachers in their effectiveness at increasing student attendance, on par with the variation in teacher effectiveness at raising student test performance. These value-added to attendance measures are as stable over time as are measures of teachers' value-added to test performance. Yet, value-added to attendance and achievement are distinct. Many teachers excel at one but not at the other. While teachers who are more effective at engagement tend to be more effective at raising achievement, this relationship is weak.

We find that teacher's ability to reduce unexcused absences contribute strongly to students' probability of completing high school and AP course taking, especially for students with lower prior attendance and lower prior achievement. Our results provide evidence of the multi-dimensional nature of teaching effectiveness, and that teachers require different skillsets to help different students succeed. A teacher who has high value-added to attendance can engage students in class and motivate the student to pursue higher academic goals. Not surprisingly, benefits from these teachers are more salient for students with low prior attendance and low prior achievement. In contrast, teachers with high value-added to achievement do little to improve graduation and reduce dropout for students prone to missing class, though they can help students pursue higher academic goals.

Our analyses build on the prior literature. While other studies have assessed teachers' contribution to attendance and find a distinction between teachers who contribute to attendance and those that contribute to achievement, ours is the first along a number of dimensions. First, we use data that identifies class-by-class unexcused absences instead of full-day absences across all middle and high school grade levels. Prior work has not look at this grade range, where unexcused absences are the most common and teachers' effects on absences likely the greatest. Moreover, prior studies have used all absences instead of distinguishing unexcused absences, which, as we demonstrate, teachers are more likely to affect. Second, we use the NBRM model, an approach that is more appropriate for dealing with count data like attendance. Finally, and most substantively, we are able to assess the longer-run effects of teacher value-added to attendance, demonstrating both the predictive validity of the measure and the importance of this dimension of teacher effectiveness for students' academic accomplishments.

Our results, particularly in conjunction with other recent papers (Gershenson, 2016; Jackson, 2018) confirm that teacher effectiveness is multi-dimensional. Effectiveness at improving student test performance does not fully capture the qualities of teachers that benefit students in the long run. Moreover, the results provide evidence that teachers' ability to engage students in class (have them show up), in particular, is an important dimension of teacher effectiveness, especially for boosting students' likelihood of graduating from high school. Finally, the importance of engaging teachers, combined with the substantial extent of unexcused class skipping, points, more broadly, to the importance of better engaging students, whether that is done by teachers or by other experiences in or out of school.

Reference:

Allensworth, Elaine, and John Easton. 2007. "What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year. Research Report." *Consortium on Chicago School Research*.

Allison, Paul D, and Richard P Waterman. 2002. "Fixed-Effect Negative Binomial Regression Models." *Sociological Methodology* 32(1): 247-265.

Atteberry, Allison, Susanna Loeb, and James Wyckoff. 2013. "Do First Impressions Matter? Improvement in Early Career Teacher Effectiveness." *NBER Working Paper* 1 (February): 51.

Backes, Ben, and Michael Hansen. 2015. "Teach For America Impact Estimates on Nontested Student Outcomes." *National Center for Analysis of Longitudinal in Education Research Working Paper*, *146*.

Balfanz, Robert, and V Byrnes. 2012. "The Importance of Being in School: A Report on Absenteeism in the Nation's Public Schools." Baltimore: Johns Hopkins University Center for Social Organization of Schools.

Balfanz, Robert, Liza Herzog, and Douglas J. Mac Iver. 2007. "Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions." *Educational Psychologist* 42 (4): 223–35.

Blazar, D., and M. A. Kraft. 2017. "Teacher and Teaching Effects on Students Attitudes and Behaviors." *Educational Evaluation and Policy Analysis* 39(1):146-70

Chamberlain, Gary E. 2013. "Predictive Effects of Teachers and Schools on Test Scores, College Attendance, and Earnings." *Proceedings of the National Academy of Sciences of the United States of America* 110 (43): 17176–82.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2007. "Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects." *Economics of Education Review* 26 (6): 673–82.

Cortes, Kalena E., Jesse Bricker, and Chris Rohlfs. 2012. "The Role of Specific Subjects in Education Production Functions: Evidence from Morning Classes in Chicago Public High Schools." *B.E. Journal of Economic Analysis and Policy* 12 (1).

Gershenson, Seth. 2016. "Linking Teacher Quality, Student Attendance, and Student Achievement." *Education Finance and Policy*, 11(2):125-49.

Glenn Ellison. 2016. "Do Schools Matter for High Math Achievement?" *American Economic Review* 106 (6): 1244–77.

Goldhaber, Dan D. 2007. "Everyone's Doing It, but What Does Teacher Testing Tell Us about Teacher Effectiveness?" *Journal of Human Resources* 42 (4): 765–94.

Goodman, Joshua. 2014. "Flaking out: Student Absences and Snow Days as Disruptions of Instruction Time." *NBER Working Paper* 20221.

Gottfried, Michael A. 2009. "Excused Versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement." *Educational Evaluation and Policy Analysis* 31 (4): 392–415.

———. 2011. "Evaluating the Relationship Between Student Attendance and Achievement in

Urban Elementary and Middle Schools: An Instrumental Variables Approach." *American Educational Research Journal* 47 (2): 434–65.

———. 2011. "The Detrimental Effects of Missing School: Evidence from Urban Siblings." *American Journal of Education* 117 (2): 147–82.

Guarino, Cassandra, Mark Reckase, Brian Stacy, and Jeffrey Wooldridge. 2015. "A Comparison of Student Growth Percentile and Value-Added Models of Teacher Performance." *Statistics and Public Policy* 2(1):1-1.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "Disruption versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics* 88 (9–10): 1721–46.

Hausman, J.A., Bronwyn H. Hall, and Z. Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica*.

Jackson, C. Kirabo. 2014 "Teacher quality at the high school level: The importance of accounting for tracks." *Journal of Labor Economics 32*(4): 645-684.

———. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." *Journal of Political Economy*. Vol. No. 22226. Cambridge, MA.

Jennings, Jennifer L., and Thomas A. DiPrete. 2010. "Teacher Effects on Social and Behavioral Skills in Early Elementary School." *Sociology of Education* 83 (2): 135–59.

Kane, Thomas J, and Douglas O Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." *National Bureau of Economic Research*, no. 14607: 51.

Kraft, Matthew A., and Sarah Grace. 2015. "Teaching for Tomorrow's Economy? Elementary Teacher Effects on Complex Tasks, Grit, and Growth Mindset." Working Paper.

Ladd, Helen F., and Lucy C. Sorensen. 2016. "Returns to Teacher Experience: Student Achievement and Motivation in Middle School." *Education Finance and Policy*, 12(2):241-79.

Long, J Scott, and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. *Stata Press Books*.

McCaffrey, D. F., J. R. Lockwood, D. Koretz, T. a. Louis, and L. Hamilton. 2004. "Models for Value-Added Modeling of Teacher Effects." *Journal of Educational and Behavioral Statistics* 29 (1): 67–101.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata, Second Edition*. Stata Press.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2): 417–58.

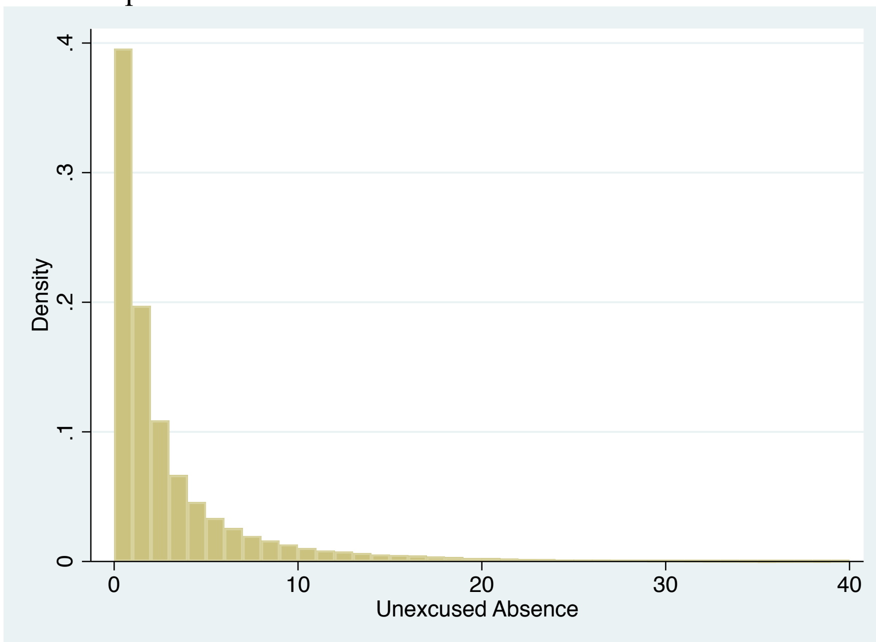
Romero, Mariajose, and Young-Sun Lee. 2007. "The Influence of Maternal and Family Risk on Chronic Absenteeism in Early Schooling." *National Center for Children in Poverty*, December.

Ruzek, E. A., T. Domina, A. M. Conley, G. J. Duncan, and S. A. Karabenick. 2015. "Using Value-Added Models to Measure Teacher Effects on Students' Motivation and Achievement." *The Journal of Early Adolescence* 35 (5–6): 852–82.

Snyder, Thomas D., and Sally A. Dillow. 2013. "Digest of Education Statistics, 2012. NCES 2014-015." *National Center for Education Statistics*.

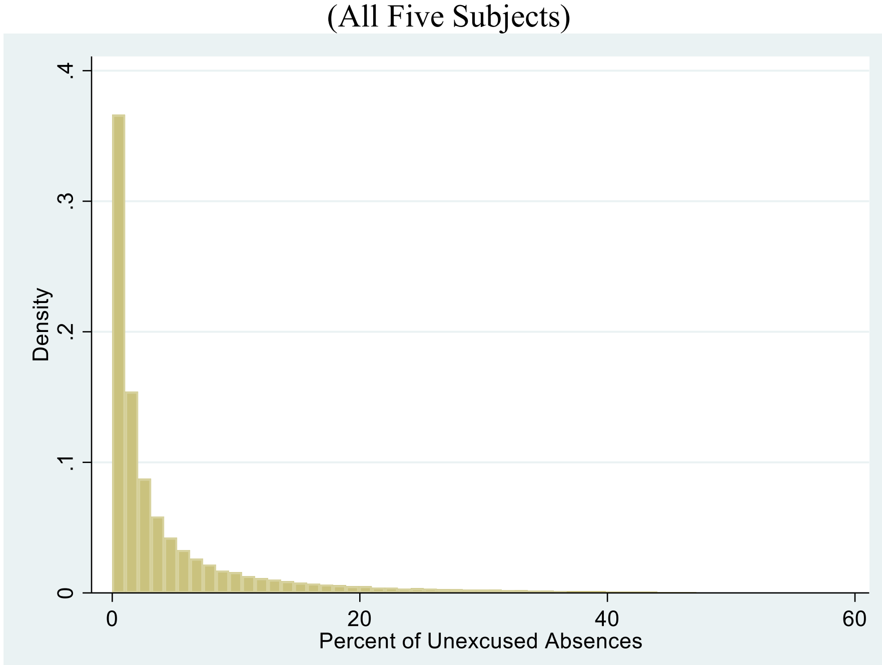
Whitney, Camille.R., and Jing Liu. (2017) "What we're missing: A descriptive analysis of part-day absenteeism in secondary school." *AERA Open* 3, no. 2: 2332858417703660.

Graph 1.1 Distribution of Unexcused Absences for Math



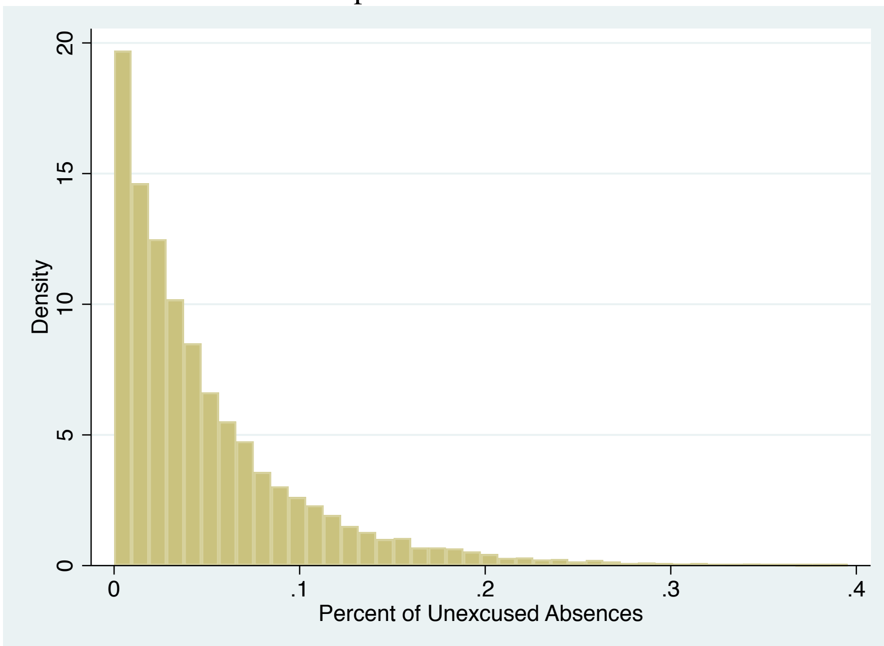
Note: This graph uses the restricted data at the student-class period-semester level. The data are truncated at 40 absences to show the bulk of the distribution.

Graph 1.2 Distribution of Percent of Unexcused Absences over Total Class Meetings (All Five Subjects)



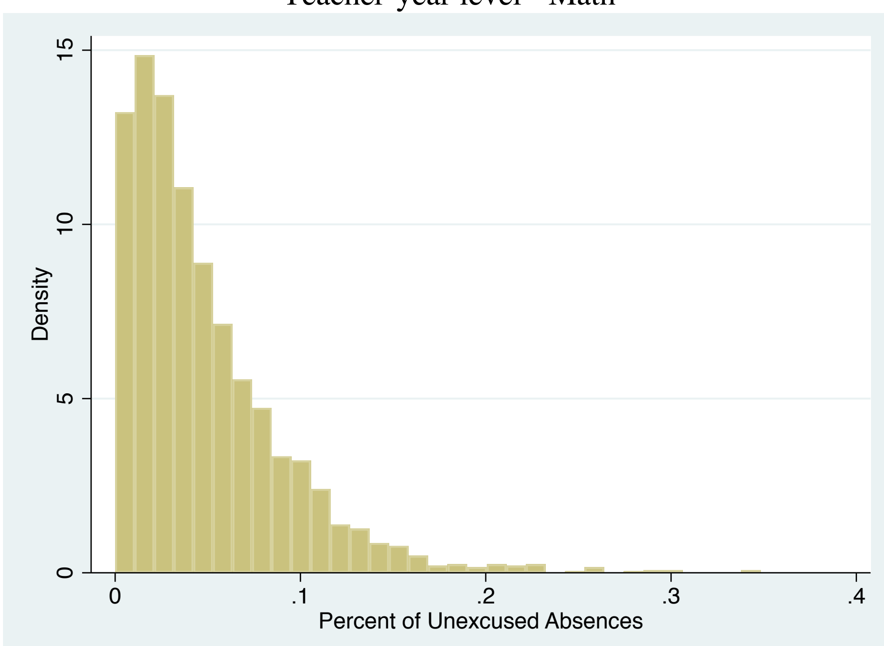
Note: This graph uses the restricted data that have all five subjects at the student-school year level.

Graph 1.3 Distribution of Percent of Unexcused Absences over Total Class Meetings
Class-period level — Math



Note: This graph uses data collapsed at the class-period Level for all math classes during school years 2002-3003 to 2012-2013.

Graph 1.4 Distribution of Percent of Unexcused Absences over Total Class Meetings
Teacher-year level - Math



Note: This graph uses data collapsed at the teacher level for all math classes during school years 2002-3003 to 2012-2013.

Graph 2 Binned Scatter Plot:
Teacher Effects on Attendance Versus. Teacher Effects on Achievement

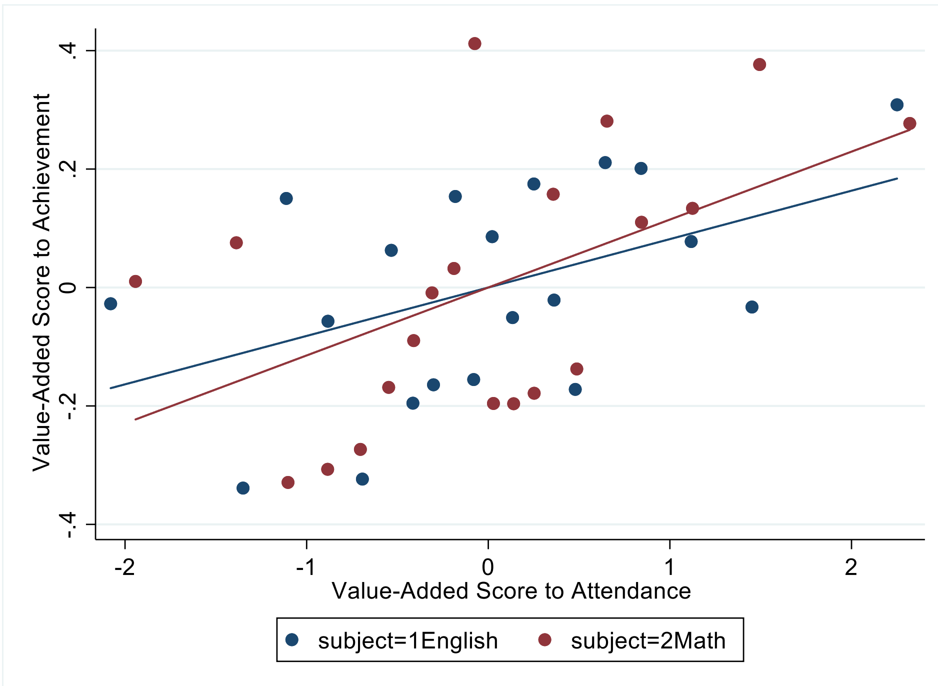


Table 1 Descriptive Statistics

	Full Sample		Analytical Sample					
Variable	All Subjects		All Subjects		Math		ELA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Student								
Female	0.484		0.486		0.483		0.484	
White	0.081		0.081		0.077		0.080	
Black	0.113		0.103		0.091		0.094	
Hispanic	0.218		0.209		0.203		0.197	
Asian	0.498		0.519		0.540		0.537	
EL	0.205		0.197		0.178		0.142	
Excused Absences	1.724	(2.821)	1.693	(2.577)	1.602	(2.502)	1.634	(2.518)
Unexcused Absences	3.765	(7.000)	3.373	(5.200)	3.040	(4.962)	2.959	(4.901)
Total Class Meetings	75.069	(17.305)	76.155	(13.918)	76.768	(14.196)	77.132	(13.494)
Math Score	0.026	(0.999)	0.058	(0.977)	0.080	(0.948)	0.072	(0.928)
ELA Score	0.008	(0.998)	0.044	(0.973)	0.080	(0.925)	0.110	(0.901)
Class-period								
White	0.081	(0.088)	0.082	(0.090)	0.080	(0.095)	0.084	(0.100)
Black	0.111	(0.140)	0.102	(0.140)	0.097	(0.151)	0.100	(0.149)
Hispanic	0.218	(0.211)	0.208	(0.208)	0.203	(0.224)	0.198	(0.215)
Asian	0.502	(0.250)	0.521	(0.249)	0.531	(0.274)	0.526	(0.254)
EL	0.189	(0.014)	0.189	(0.036)	0.188	(0.037)	0.188	(0.036)
Excused Absences	1.731	(1.090)	1.708	(1.061)	1.689	(1.176)	1.725	(1.163)
Unexcused Absences	3.785	(3.661)	3.970	(3.667)	3.923	(4.249)	3.943	(3.992)
Total Class Meetings	75.348	(16.151)	74.209	(16.983)	74.417	(17.574)	74.419	(17.459)
Math Score	-0.006	(0.670)	0.010	(0.684)	0.037	(0.754)	-0.004	(0.701)
ELA Score	-0.025	(0.726)	-0.009	(0.727)	0.009	(0.755)	-0.026	(0.773)
School								

White	0.084	(0.069)	0.084	(0.070)	0.083	(0.068)	0.081	(0.070)
Black	0.111	(0.093)	0.105	(0.094)	0.102	(0.093)	0.102	(0.093)
Hispanic	0.214	(0.174)	0.206	(0.171)	0.204	(0.171)	0.203	(0.171)
Asian	0.503	(0.201)	0.517	(0.200)	0.523	(0.199)	0.525	(0.199)
EL	0.204	(0.149)	0.200	(0.150)	0.197	(0.151)	0.198	(0.149)
Excused Absences	1.745	(0.767)	1.644	(0.711)	1.640	(0.701)	1.633	(0.695)
Unexcused Absences	3.677	(2.574)	3.097	(1.878)	3.003	(1.694)	3.061	(1.672)
Total Class Meetings	75.436	(15.402)	76.673	(12.179)	77.207	(11.864)	76.924	(11.927)
Math Score	0.002	(0.499)	0.032	(0.474)	0.016	(0.401)	-0.002	(0.368)
ELA Score	-0.013	(0.528)	0.018	(0.509)	0.008	(0.417)	-0.009	(0.377)
Observations								
Student by Year	230,686		184,976		136,540		124,800	
Teacher by Year	11,372		8,893		2,510		2,606	
School by Year	367		367		367		367	

Note: Data are for all students in 6th to 11th grades from school year 2003-2004 through 2012-2013 as we use school year 2002-2003 to generate prior achievement and attendance. Characteristics are calculated using the final matched data sets at student-year level. "All subjects" include math, ELA, science, social studies and foreign languages. At student level, absences and total class meetings are averages across all class-periods taken in the corresponding subject in a school year. To construct the analytical sample, we drop observations when a student has more than one teacher in a subject for the entire school year, is absent from more than 50 percent of his classes, has less than ten valid attendance marks in a class per semester, and classes have fewer than five students.

Table 2 Characteristics Predicting Unexcused Class Absence Rate

	All Subjects		Math		ELA	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.003**	-0.002**	-0.002**	-0.002**	-0.002**	-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
White	0.020**	0.021**	0.021**	0.022**	0.019**	0.021**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black	0.074**	0.063**	0.073**	0.063**	0.070**	0.063**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	0.045**	0.039**	0.045**	0.039**	0.042**	0.037**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Other	0.025**	0.023**	0.026**	0.023**	0.025**	0.023**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
English Language Learner	0.014**	0.010**	0.014**	0.011**	0.014**	0.012**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Grade 8	0.002**	0.002**	0.003**	0.002**	0.003**	0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Grade 9	0.031**	0.010**	0.032**	0.015**	0.034**	-0.003
	(0.000)	(0.002)	(0.001)	(0.005)	(0.001)	(0.005)
Grade 10	0.032**	0.010**	0.032**	0.016**	0.032**	-0.004
	(0.000)	(0.002)	(0.001)	(0.005)	(0.001)	(0.005)
Grade 11	0.030**	0.010**	0.032**	0.017**	0.029**	-0.006
	(0.000)	(0.002)	(0.001)	(0.005)	(0.001)	(0.005)
2nd Period	-0.009**	-0.009**	-0.008**	-0.008**	-0.008**	-0.008**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
3rd Period	-0.012**	-0.012**	-0.011**	-0.010**	-0.010**	-0.010**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
4th Period	-0.010**	-0.011**	-0.007**	-0.009**	-0.010**	-0.010**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
5th Period	-0.010**	-0.010**	-0.009**	-0.009**	-0.008**	-0.008**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
6th Period	-0.007**	-0.007**	-0.007**	-0.006**	-0.007**	-0.006**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
7th Period	-0.004**	-0.002**	-0.001	-0.000	-0.001	-0.002*
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Other Periods	-0.022**	-0.004**	-0.022**	-0.002	-0.012**	-0.008**
	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
Math	0.001**	0.002**				
	(0.000)	(0.000)				

Science	0.001**	0.003**				
	(0.000)	(0.000)				
Social Studies	0.001+	0.002**				
	(0.000)	(0.000)				
Foreign language	-0.003**	0.001**				
	(0.000)	(0.000)				
School by Year FE		X		X		X
Observations	1,197,741	1,197,741	262,993	262,993	253,235	253,235

Note: Robust standard errors in brackets are adjusted for clustering at class-period level. ** p<0.01, * p<0.05, + p<0.10. The dependent variable is unexcused absence rate at class-period level. All subjects include math, ELA, science, social studies and foreign languages. The reference group for the race/ethnicity variable is Asian students. The reference group for the period variable is the 1st period. The reference group for the subject variable is English classes.

Table 3 Magnitude of Teacher and Teacher-by-Year Effects on Student Absences

		Standard Deviation	Incidence Rate Ratio
Teacher	Math	0.366	1.443
	ELA	0.433	1.541
	Science	0.422	1.525
	Social Studies	0.402	1.495
	Foreign Languages	0.403	1.496
Teacher by Year	Math	0.447	1.564
	ELA	0.478	1.612
	Science	0.479	1.615
	Social Studies	0.467	1.595
	Foreign Languages	0.409	1.505

Note: Standard deviations are directly estimated from two-level Negative Binomial models. Incidence rate ratio is calculated using $\exp(SD)$.

Table 4.1 Transition Matrix: VA to Attendance

Initial Quintile		Q1	Q2	Q3	Q4	Q5	Row
		Quintile of Future Performance on Attendance					
Q1	n	35	18	13	12	1	79
	(row %)	(44.30)	(22.78)	(16.46)	(15.19)	(1.27)	(100.00)
Q2	n	24	12	20	13	8	77
	(row %)	(31.17)	(15.58)	(25.97)	(16.88)	(10.39)	(100.00)
Q3	n	10	22	18	19	9	78
	(row %)	(12.82)	(28.21)	(23.08)	(24.36)	(11.54)	(100.00)
Q4	n	6	19	20	12	20	77
	(row %)	(7.79)	(24.68)	(25.97)	(15.58)	(25.97)	(100.00)
Q5	n	4	6	7	21	39	77
	(row %)	(5.19)	(7.79)	(9.09)	(27.27)	(50.65)	(100.00)
Column Total		79	77	78	77	77	388

Note (same for Table 4.2): Only using teachers who have at least five years' observations in our sample. Bottom quintiles represent those who are least effective in reducing unexcused absences. We combine math and English teachers together, though we calculate their quintiles by subject.

Table 4.2 Transition Matrix: VA to Achievement

Initial Quintile		Q1	Q2	Q3	Q4	Q5	Row
		Quintile of Future Performance on Achievement					
Q1	n	36	19	11	11	2	79
	(row %)	(45.57)	(24.05)	(13.92)	(13.92)	(2.53)	(100.00)
Q2	n	24	21	18	12	2	77
	(row %)	(31.17)	(27.27)	(23.38)	(15.58)	(2.60)	(100.00)
Q3	n	8	14	25	18	13	78
	(row %)	(10.26)	(17.95)	(32.05)	(23.08)	(16.67)	(100.00)
Q4	n	7	17	18	16	19	77
	(row %)	(9.09)	(22.08)	(23.38)	(20.78)	(24.68)	(100.00)
Q5	n	4	6	6	20	41	77
	(row %)	(5.19)	(7.79)	(7.79)	(25.97)	(53.25)	(100.00)
Column Total		79	77	78	77	77	388

Table 5 Adjusted R-Squared Using Early Year VA to Predict Future Performance

Early Year VA Predictor(s)	*Outcome (Attendance)* VA in Y3	VA in Y4	VA in Y5	Mean(VA_{Y3-5})
Math				
Math VA in Y1 Only	0.222	0.205	0.076	0.226
Math VA in Y2 Only	0.260	0.312	0.172	0.355
Math VA in Y1 & Y2	0.319	0.349	0.174	0.397
ELA				
ELA VA in Y1 Only	0.282	0.108	0.078	0.216
ELA VA in Y2 Only	0.309	0.140	0.088	0.251
ELA VA in Y1 & Y2	0.222	0.205	0.076	0.226
	Outcome (Achievement)			
Early Year VA Predictor(s)	VA in Y3	VA in Y4	VA in Y5	Mean(VA_{Y3-5})
Math				
Math VA in Y1 Only	0.216	0.260	0.098	0.280
Math VA in Y2 Only	0.315	0.176	0.094	0.282
Math VA in Y1 & Y2	0.352	0.289	0.123	0.363
ELA				
ELA VA in Y1 Only	0.112	0.130	0.066	0.209
ELA VA in Y2 Only	0.148	0.148	0.088	0.265
ELA VA in Y1 & Y2	0.216	0.260	0.098	0.280

Note: Only using teachers who have at least five years' observations in our sample. All entries are adjusted R-squared.

Table 6. Effects of Out of Sample Teacher Effects on Current Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Test Scores			Unexcused Absence Rate	
Test Score VA	0.08161**		0.08160**	-0.00046**		0.00011
	(0.00154)		(0.00155)	(0.00013)		(0.00013)
Attendance VA		0.00946**	0.00013		-0.00790**	-0.00791**
		(0.00183)	(0.00182)		(0.00016)	(0.00016)
Observations	223623	223623	223623	223623	223623	223623
Adjusted R2	0.657	0.653	0.657	0.428	0.434	0.434

Note: Each column reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Dependent variables are current test scores and unexcused absence rates. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; subject fixed effects; and school fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Table 7. Effects of Out of Sample Teacher Effects on Long-Term Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Graduation			Dropout Before 12th Grade	
Test Score VA	0.00117		0.00055	-0.00050		-0.00024
	(0.00109)		(0.00111)	(0.00076)		(0.00077)
Attendance VA		0.00710**	0.00702**		-0.00293**	-0.00289**
		(0.00126)	(0.00128)		(0.00086)	(0.00088)
Observations	197639	197639	197639	197639	197639	197639
Adjusted R2	0.208	0.208	0.208	0.107	0.107	0.107
		Number of AP Courses			Credits of AP Courses	
Test Score VA	0.02272**		0.02168**	0.11314**		0.10771**
	(0.00239)		(0.00239)	(0.01189)		(0.01189)
Attendance VA		0.01470**	0.01172**		0.07561**	0.06082**
		(0.00310)	(0.00310)		(0.01541)	(0.01542)
Observations	197639	197639	197639	197639	197639	197639
Adjusted R2	0.306	0.306	0.306	0.305	0.305	0.305

Note: Each column, within panels, reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both the number and credits earned for AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; subject fixed effects; and school fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Table 8. Effects of Out of Sample Teacher Effects on Long-Term Outcomes: Nonlinearity

	(1)	(2)	(3)	(4)
	Graduation	Dropout Before 12th Grade	Number of AP Courses	Earned Credits of AP Courses
Test Score VA	0.00033	-0.00015	0.02071**	0.10322**
	(0.00112)	(0.00078)	(0.00235)	(0.01169)
Test Score VA_2	-0.00010	-0.00023	0.00366**	0.01697**
	(0.00062)	(0.00045)	(0.00116)	(0.00576)
Attendance VA	0.00911**	-0.00320**	0.01183**	0.06135**
	(0.00133)	(0.00090)	(0.00327)	(0.01628)
Attendance VA_2	-0.00231**	0.00036	-0.00040	-0.00186
	(0.00049)	(0.00036)	(0.00099)	(0.00494)
Observations	197639	197639	197639	197639
Adjusted R2	0.208	0.107	0.306	0.305

Note: Each column reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both the number and credits of AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; subject fixed effects; and school fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** $p<0.01$, * $p<0.05$, + $p<0.10$.

Table 9. Effects of Out of Sample Teacher Effects on Long-Term Outcomes: By Tertiles of Prior Attendance/Achievement

	Graduation	Dropout Before 12th Grade	Number of AP Courses	Earned Credits of AP Courses	Graduation	Dropout Before 12th Grade	Number of AP Courses	Credits of AP Courses
Panel A. Tertiles of prior attendance								
	Bottom tertile of attendance				Top tertile of attendance			
Test Score VA	0.00256	-0.00126	0.01260**	0.06572**	-0.00002	0.00020	0.03067**	0.15146**
	(0.00225)	(0.00171)	(0.00276)	(0.01355)	(0.00141)	(0.00078)	(0.00511)	(0.02555)
Attendance VA	0.01130**	-0.00468*	0.01433**	0.07091**	0.00222	-0.00056	-0.00636	-0.02842
	(0.00299)	(0.00220)	(0.00392)	(0.01940)	(0.00145)	(0.00082)	(0.00595)	(0.02968)
Observations	64366	64366	64366	64366	67134	67134	67134	67134
Adjusted R-squared	0.187	0.114	0.215	0.214	0.073	0.032	0.306	0.305
Panel B. Tertiles of prior math scores								
	Bottom tertile of math scores				Top tertile of math scores			
Test Score VA	0.00566**	-0.00134	0.00841**	0.04274**	0.00099	-0.00072	0.03153**	0.15885**
	(0.00218)	(0.00157)	(0.00193)	(0.00942)	(0.00144)	(0.00092)	(0.00600)	(0.02996)
Attendance VA	0.01365**	-0.00505*	0.00522*	0.02523*	0.00281+	-0.00175+	-0.01441*	-0.06707+
	(0.00291)	(0.00210)	(0.00254)	(0.01241)	(0.00153)	(0.00098)	(0.00702)	(0.03503)
Observations	62434	62434	62434	62434	68817	68817	68817	68817
Adjusted R-squared	0.198	0.120	0.161	0.155	0.093	0.051	0.277	0.277

Note: Each column, within panels, reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both number and credits of AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Appendix A: List of covariates included in models

Prior math test score (standardized)
Prior ELA test score (standardized)
Prior absence rate in math
Prior absence rate in ELA
Black
Hispanic
Asian
Female
English learner status
Special education status
Gifted education status
Prior suspensions
Current math test
Prior math test
Class average prior math test score
Class average prior reading test score
Class average prior absence rate
Class average prior suspensions
Class percentage black
Class percentage Hispanic
Class percentage Asian
Class percentage English learners
Class periods
School percentage black
School percentage Hispanic
School percentage Asian
School percentage English learners
School average prior absence rate
School average prior suspensions

Appendix B: Regression results of estimating value-added to attendance for math teachers

	Coefficient		Standard Error
% Any absence_math_lag	1.500	**	0.103
% Any absence_ELA_lag	1.734	**	0.098
Any absence_lag	4.982	**	0.102
Test score_math_lag (standardized)	-0.170	**	0.006
Test score_ELA_lag (standardized)	-0.039	**	0.006
Suspension days_lag	0.020	**	0.004
Black	0.159	**	0.014
Hispanics	0.099	**	0.011
Asian	-0.356	**	0.009
Female	-0.067	**	0.006
Special Education	-0.113	**	0.014
Gifted	-0.027	**	0.009
EL	0.046	**	0.011
Class black	0.129	*	0.064
Class Hispanics	-0.059		0.053
Class Asian	-0.087	+	0.047
Class EL	3.424	**	0.763
Class math test score_lag	-0.244	**	0.018
Class ELA test score_lag	-0.020		0.017
Class absence_lag	-0.083		0.148
Class suspension_lag	-0.005		0.015
School black	-0.553	*	0.232
School Hispanics	-0.508	**	0.173
School Asian	-0.388	*	0.160
School EL	0.335	**	0.104
School absence_lag	2.776	**	0.441
School suspension_lag	-0.188	**	0.054
grade = 8	0.018		0.017
grade = 9	0.586	**	0.057
grade = 10	0.294	**	0.072
grade = 11	0.378	**	0.075
Period = 2	-0.304	**	0.012
Period = 3	-0.345	**	0.012
Period = 4	-0.313	**	0.012
Period = 5	-0.268	**	0.013
Period = 6	-0.190	**	0.012
Period = 7	0.010		0.016

Period = 8	-0.048		0.059
Constant	-4.631	**	0.195
ln(total)	1.000		(exposure)
/lnalpha	-0.217		0.006
var(_cons)	0.200		0.007

Appendix C: Estimating value-added to attendance using NBRM and controlling for school fixed effects

Table C1. Magnitude of Teacher Effects on Class Unexcused Absences

		Standard Deviation	Incidence Rate Ratio
Teacher	Math	0.280	1.324
	ELA	0.352	1.422
	Science	0.322	1.380
	Social Studies	0.310	1.363
	Foreign Languages	0.281	1.325
Teacher by Year	Math	0.361	1.435
	ELA	0.393	1.482
	Science	0.377	1.458
	Social Studies	0.382	1.465
	Foreign Languages	0.330	1.392

Table C2. Transition Matrix: VA to Attendance

		Quintile of Future Performance on Attendance					
Initial Quintile		Q1	Q2	Q3	Q4	Q5	Row
Q1	n	27	15	21	15	1	79
	(row %)	(34.18)	(18.99)	(26.58)	(18.99)	(1.27)	(100)
Q2	n	24	17	16	11	9	77
	(row %)	(31.17)	(22.08)	(20.78)	(14.29)	(11.69)	(100)
Q3	n	15	30	12	12	9	78
	(row %)	(19.23)	(38.46)	(15.38)	(15.38)	(11.54)	(100)
Q4	n	7	7	19	25	19	77
	(row %)	(9.09)	(9.09)	(24.68)	(32.47)	(24.68)	(100)
Q5	n	6	8	10	14	39	77
	(row %)	(7.79)	(10.39)	(12.99)	(18.18)	(50.65)	(100)
Column Total		79	77	78	77	77	**388**

Table C3. Adjusted R-Squared Using Early Year VA to Predict Future Performance

Early Year VA Predictor(s)	Outcome (Attendance)			
	VA in Y3	VA in Y4	VA in Y5	Mean(VA_{Y3-5})
Math				
Math VA in Y1 Only	0.030	0.064	0.000	0.065
Math VA in Y2 Only	0.028	0.102	0.023	0.083
Math VA in Y1 & Y2	0.043	0.129	0.019	0.115
ELA				
ELA VA in Y1 Only	0.037	0.037	0.002	0.040
ELA VA in Y2 Only	0.096	0.025	0.001	0.060
ELA VA in Y1 & Y2	0.102	0.044	0.000	0.075

Appendix D: Joint estimates of value-added to attendance and value-added to achievement

The purpose of this model is to get the true correlation between the two types of value-added scores through running this joint model. We cannot run a negative binomial model in this case, thus we reconstruct the unexcused absence outcome to unexcused absence rate (standardized by year and grade). Then we create a variable called "Outcome" and two indicator variables, "TestScore" and "Absence." "Outcome" takes the values of test scores when the dummy variable "TestScore" equals 1, and takes the values of unexcused absence rate when the dummy variable "Absence" equals 1.

We estimate the model is below.
Level 1:

$$Outcome_{ijt} = \beta_{1j}\left(TestScore_{ijt}\right) + \beta_{2j}\left(Absence_{ijt}\right) \\ +\left(TestScore_{ijt}\right)X'_{ijt}\emptyset_T + \left(Absence_{ijt}\right)X'_{ijt}\emptyset_A + e_{ijt}$$

Where e_{ijt}~$\mathrm{N}(0, \sigma^2)$

Level 2:

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$
$$\beta_{2j} = \gamma_{20} + \mu_{2j}$$

Where $\begin{pmatrix}\mu_{1j} \\ \mu_{2j}\end{pmatrix} \sim N\left(\begin{matrix}0 \\ 0\end{matrix}, \begin{pmatrix}\tau_1 & \tau_{2,1} \\ \tau_{1,2} & \tau_2\end{pmatrix}\right)$

We constrain the above model to have no intercept, so that both of the indicator variables will contain a random teacher effect. The random teacher effects are assumed to have a mean of zero and a variance to be estimated. We interact all of the controls with these two indicators so that these controls can have differential effects for test scores and absences.

Appendix E: Estimating value-added to attendance using OLS

As we describe in the Methods section, students can have different "exposure" times to a teacher during a school year, thus we cannot directly do a logarithm transformation of the raw counts of absences as the dependent variable in the OLS model. Instead, we calculate the rate of unexcused absences overall total class meetings a student can have with a teacher.

As shown below, value-added to attendance using OLS is similarly stable compared with those from NBRM (see Table E1 and E2 for results in keeping with Tables 4 and 5). However, they do not show as consistent relationships with student short- and long-run outcomes. Table E3 (similar to Table 6) shows that value-added to attendance has a negative impact students' unexcused absence rate which is consistent with the NBRM results, but it also has a negative effect on student test scores, which is not consistent with the NBRM results. Similarly, in Table E4 (mirroring Table 7), it positively affects student graduation but negatively affects AP course taking. Tables E5 and E6 shows results that are quite similar to Tables 8 and 9.

To assess whether the inconsistencies in the two approaches stem from predictable shortcomings in the OLS approach, we remove teachers who are in the top quartile of having zero-absence students (approximately 35% of zero-absence students in total students a teacher has) from the analysis. Because OLS estimates a linear relationship, these students could have negative predicted absenteeism so that the residual will be positive (negative on attendance). We then replicate Table E4-a, and show the corresponding results in Table E4-b. We find that the results from the OLS and NBRM models largely converge when we remove these teachers.

Table E1. Transition Matrix: VA to Attendance

		Quintile of Future Performance on Attendance					
Initial Quintile		Q1	Q2	Q3	Q4	Q5	Row
Q1	n	38	21	9	10	2	80
	(row %)	(47.50)	(26.25)	(11.25)	(12.50)	(2.50)	(100)
Q2	n	17	26	22	8	7	80
	(row %)	(21.25)	(32.50)	(27.50)	(10.00)	(8.75)	(100)
Q3	n	8	12	23	26	10	79
	(row %)	(10.13)	(15.19)	(29.11)	(32.91)	(12.66)	(100)
Q4	n	13	17	14	17	19	80
	(row %)	(16.25)	(21.25)	(17.50)	(21.25)	(23.75)	(100)
Q5	n	4	4	11	19	41	79
	(row %)	(5.06)	(5.06)	(13.92)	(24.05)	(51.90)	(100)
Column Total		79	77	78	77	77	**388**

Table E2. Adjusted R-Squared Using Early Year VA to Predict Future Performance

	Outcome (Attendance)			
Early Year VA Predictor(s)	VA in Y3	VA in Y4	VA in Y5	Mean(VA_{Y3-5})
Math				
Math VA in Y1 Only	0.108	0.053	0.076	0.138
Math VA in Y2 Only	0.096	0.104	0.088	0.181
Math VA in Y1 & Y2	0.158	0.124	0.126	0.253
ELA				
ELA VA in Y1 Only	0.129	0.084	0.013	0.121
ELA VA in Y2 Only	0.359	0.265	0.125	0.361
ELA VA in Y1 & Y2	0.368	0.268	0.121	0.368

Table E3. Effects of Out of Sample Teacher Effects on Current Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Test Scores			Unexcused Absence Rate	
Test Score VA	0.10866**		0.10969**	-0.00071**		-0.00050**
	(0.00181)		(0.00182)	(0.00015)		(0.00015)
Attendance VA		-0.01439**	-0.01942**		-0.00404**	-0.00402**
		(0.00166)	(0.00165)		(0.00017)	(0.00017)
Observations	223623	223623	223623	223623	223623	223623
Adjusted R-squared	0.659	0.653	0.659	0.428	0.430	0.430

Note: Each column reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Dependent variables are current test scores and unexcused absence rates. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year, subject, and school fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Table E4-a. Effects of Out of Sample Teacher Effects on Long-Term Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Graduation			Dropout Before 12th Grade	
Test Score VA	0.00100		0.00086	-0.00028		-0.00025
	(0.00129)		(0.00129)	(0.00090)		(0.00090)
Attendance VA		0.00299*	0.00295*		-0.00071	-0.00070
		(0.00127)	(0.00128)		(0.00089)	(0.00089)
Observations	197639	197639	197639	197639	197639	197639
Adjusted R2	0.208	0.208	0.208	0.107	0.107	0.107
		Number of AP Courses			Credits of AP Courses	
Test Score VA	0.02490**		0.02544**	0.12334**		0.12607**
	(0.00306)		(0.00306)	(0.01522)		(0.01523)
Attendance VA		-0.01048**	-0.01151**		-0.05314**	-0.05824**
		(0.00239)	(0.00239)		(0.01188)	(0.01187)
Observations	197639	197639	197639	197639	197639	197639
Adjusted R2	0.306	0.306	0.306	0.305	0.305	0.305

Note: Each column, within panels, reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both the number and credits earned for AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Table E4-b. Effects of Out of Sample Teacher Effects on Long-Term Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
		Graduation			Dropout Before 12th Grade	
Test Score VA	0.00189		0.00081	-0.00139		-0.00094
	(0.00166)		(0.00167)	(0.00114)		(0.00115)
Attendance VA		0.01200**	0.01191**		-0.00515**	-0.00505**
		(0.00182)	(0.00183)		(0.00119)	(0.00121)
Observations	144415	144415	144415	144415	144415	144415
Adjusted R2	0.219	0.219	0.219	0.116	0.116	0.116
		Number of AP Courses			Credits of AP Courses	
Test Score VA	0.01521**		0.01497**	0.07551**		0.07388**
	(0.00328)		(0.00329)	(0.01629)		(0.01630)
Attendance VA		0.00442	0.00269		0.02656	0.01802
		(0.00381)	(0.00381)		(0.01892)	(0.01893)
Observations	144415	144415	144415	144415	144415	144415
Adjusted R2	0.281	0.281	0.281	0.280	0.280	0.280

Note: The data used are the same as Table E4-A except excluding teachers who are in the top quartile of having zero-absence students. Each column, within panels, reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both the number and credits earned for AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Table E5. Effects of Out of Sample Teacher Effects on Long-Term Outcomes: Nonlinearity

	(1)	(2)	(3)	(4)
	Graduation	Dropout Before 12th Grade	Number of AP Courses	Earned Credits of AP Courses
Test Score VA	0.00333+	-0.00096	0.03429**	0.16986**
	(0.00172)	(0.00120)	(0.00365)	(0.01816)
Test Score VA_Squared	-0.00028	-0.00053	0.02567**	0.12540**
	(0.00149)	(0.00107)	(0.00273)	(0.01357)
Attendance VA	0.00702**	-0.00545**	0.03918**	0.19712**
	(0.00111)	(0.00077)	(0.00277)	(0.01379)
Attendance VA_Squared	-0.00269**	0.00053+	-0.00357**	-0.01764**
	(0.00044)	(0.00031)	(0.00094)	(0.00471)
Observations	197639	197639	197639	197639
Adjusted R2	0.205	0.104	0.286	0.285

Note: Each column reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both the number and credits of AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** $p<0.01$, * $p<0.05$, + $p<0.10$.

Table E6. Effects of Out of Sample Teacher Effects on Long-Term Outcomes: By Tertiles of Prior Attendance/Achievement

	Graduation	Dropout Before 12th Grade	Number of AP Courses	Earned Credits of AP Courses	Graduation	Dropout Before 12th Grade	Number of AP Courses	Credits of AP Courses
Panel A. Tertiles of prior attendance								
	Bottom tertile of attendance				Top tertile of attendance			
Test Score VA	0.00632+	-0.00254	0.02506**	0.12826**	0.00248	-0.00016	0.02957**	0.21589**
	(0.00342)	(0.00261)	(0.00422)	(0.02071)	(0.00219)	(0.00121)	(0.00682)	(0.04023)
Attendance VA	0.00848**	-0.00540**	0.03340**	0.16778**	-0.00076	-0.00178*	0.00012	0.02019
	(0.00237)	(0.00178)	(0.00361)	(0.01786)	(0.00134)	(0.00075)	(0.00367)	(0.02494)
Observations	64366	64366	64366	64366	67134	67134	67134	67134
Adjusted R-squared	0.184	0.111	0.195	0.193	0.068	0.028	0.286	0.286
Panel B. Tertiles of prior math scores								
	Bottom tertile of math scores				Top tertile of math scores			
Test Score VA	0.01378**	-0.00422+	0.01968**	0.09599**	0.00324	-0.00231+	0.04323**	0.21842**
	(0.00331)	(0.00241)	(0.00302)	(0.01472)	(0.00222)	(0.00140)	(0.00943)	(0.04709)
Attendance VA	0.01274**	-0.00628**	0.00582**	0.02841**	-0.00039	-0.00370**	0.00325	0.01889
	(0.00236)	(0.00175)	(0.00197)	(0.00961)	(0.00139)	(0.00087)	(0.00570)	(0.02848)
Observations	62434	62434	62434	62434	68817	68817	68817	68817
Adjusted R-squared	0.195	0.117	0.145	0.140	0.089	0.048	0.258	0.258

Note: Each column, within panels, reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using the stacked sample that pools together both math and ELA for 7th to 11th graders. Both number and credits of AP courses only include those taken in 12th grade to avoid mechanical endogeneity. All columns control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Appendix F: Selection on observables

One threat to our identification is that students are selected to teachers based on observed student characteristics. The assumption is that conditional on the controls in our specification, there is no systematic sorting of students to teachers. To test whether this is true, following Chetty et al. (2014) and Jackson (2018), we first use twice lagged student characteristics to predict all the long-term outcomes, which effectively limits our sample to students who have twice-lagged controls and only 8th to 11th graders. The student characteristics used here include test scores and absence rates for both math and ELA classes, days of suspension, race, gender, special education status, gifted status, and EL status. Using predicted outcomes and conditional on all student, class, and school characteristics excluding those used in the prediction, we should not observe any significant association between the estimated teacher value-added (leave-year-out estimates) and the predicted outcomes.

Table F1 presents the results. Columns (1) and (2) show results for using actual graduation, dropout, and AP course taking as outcomes. The results are slightly different from Table 7 because here we are only using 8th to 11th grades in the sample. We still find similar magnitudes and significance of value-added to attendance on all the long-term outcomes. In columns (3) and (4), we use the predicted outcomes as dependent variables. The sample sizes are smaller than columns (1) and (2) because of the reason we described above. We find that the significance of value-added to attendance disappear for predicted graduation and dropout. Although they are significant for predicted number of AP courses and earned AP credits, the magnitudes are so small and nearly negligible, suggesting minimum selection in our model. For value-added to achievement, although we observe significant coefficients for predicted graduation and dropout, but the directions are opposite of what we would sort of sorting. Similar to value-added to attendance, the coefficients of value-added to achievement on predicted AP course are very small and only marginally significant. Our results suggest that our strategy largely eliminate selection on observables.

Table F1. Robustness Check of Selection on Observables

	(1)	(2)	(3)	(4)
	Graduation	Dropout Before 12th Grade	Predicted: Graduation	Predicted: Dropout Before 12th Grade
Test Score VA	0.00268+	-0.00029	-0.00146**	0.00034*
	(0.00156)	(0.00104)	(0.00037)	(0.00016)
Attendance VA	0.00351**	-0.00285**	-0.00027	-0.00002
	(0.00091)	(0.00062)	(0.00021)	(0.00009)
Observations	166136	166136	129252	129252
	Number of AP Courses	Earned Credits of AP Courses	Predicted: Number of AP Courses	Predicted: Earned Credits of AP Courses
Test Score VA	0.02556**	0.12495**	0.00225*	0.01182*
	(0.00354)	(0.01762)	(0.00103)	(0.00514)

Attendance VA	0.04085**	0.20668**	0.00279**	0.01416**
	(0.00229)	(0.01142)	(0.00059)	(0.00293)
Observations	166136	166136	129252	129252

Note: Columns (3) - (4) use predicted graduation, dropout, number of AP courses, and earned credits of AP courses as outcomes. The prediction of graduation and dropout is conducted for each of 8^{th}-11^{th} grade separately using a Logit model. Predictors include twice-lagged math and English scores, test types, absence rates, and suspension, race, gender, special education status, gifted status, and EL status. Each column reports coefficients from an OLS regression, with standard errors clustered at both student and teacher level to account for correlation between observations. The columns are estimated using data pooling across 8th-11th grades and those who have twice-lagged controls. All columns control for the baseline student, class, and school level characteristics but excludes those used in the prediction, which include one-year lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade; year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.

Appendix G: Selection on unobservables

Following Jackson (2018), we test selection on unobservables based on two distinct sources of variation. The first strategy relies on school-by-cohort fixed effects. Since Jackson (2018) only uses 9th graders, he uses school-by-year fixed effects. Here we modify his approach by using school-by-cohort fixed effects because the selection of students to teachers most likely happens within school-cohort, which is most susceptible to selection on unobservables. This approach should be robust to any school-level policy and shocks. The second strategy uses a Two-Stage Least Square estimator, relying on variation induced by average estimated teacher value-added scores across cohorts within a school. This Instrumental Variable approach is robust to student selection to teachers within a school, but is susceptible to school polices or changes. If these two distinct identification strategies provide similar results, then we have extra evidence to say that our estimation strategy is not biased due to unobservables.

Table G1 present the results using the above two identification strategies. Columns (1) and (2) are from models with school-cohort fixed effects, and (3) and (4) are from models using across cohorts within school value-added as an instrumental variable. The overall magnitude and significance is remarkably consistent with Table 7, especially for value-added to attendance.

Table G1. Robustness Check of Selection on Unobservables

	(1)	(2)	(3)	(4)
	OLS with School-Cohort Fixed Effects		2SLS using Average Teacher Quality in the School-Cohort as an Instrument	
	Graduation	Dropout Before 12th Grade	Graduation	Dropout Before 12th Grade
Test Score VA	-0.00028	0.00015	0.00479	-0.00177
	(0.00152)	(0.00105)	(0.00449)	(0.00302)
Attendance VA	0.00684**	-0.00298**	0.00342*	-0.00684**
	(0.00101)	(0.00068)	(0.00133)	(0.00091)
Observations	197639	197639	197639	197639
	Number of AP Courses	Earned Credits of AP Courses	Number of AP Courses	Earned Credits of AP Courses
Test Score VA	0.02505**	0.12427**	0.01163	0.04948
	(0.00311)	(0.01550)	(0.00932)	(0.04641)
Attendance VA	0.02681**	0.13394**	0.05494**	0.27459**
	(0.00215)	(0.01072)	(0.00316)	(0.01575)
Observations	197639	197639	197639	197639

Note: Columns (1) and (2) report results from a school-cohort fixed effects model. Columns (3) and (4) report results from a Two-Stage Least Squares model by using the average teacher value-added across cohorts within schools as instruments. Standard errors are clustered at both student and teacher level to account for correlation between observations. The columns are estimated using data pooling across 7th-11th grades. All models control for the baseline student, class, and school level characteristics, which include lagged math and English scores, absence rates, suspension, and demographic composition; tests students took in both previous and current year interacted with grade;

year fixed effects; and subject fixed effects. Value-added scores are "leave-year-out" estimates standardized using "true" standard deviations of teacher effects estimated using all years of data. ** p<0.01, * p<0.05, + p<0.10.