ONLINE APPENDIX
Does Helping John Help Sue? Evidence of Spillovers in Education
Isaac M. Opper

A. Data, Context, and Teacher Value-Added Estimation

Misclassification In addition to the sample restrictions, I correct information that appears to be a misclassification. In particular, I code as missing elementary school teachers who are initially assigned to more than 50 students or less than 10 students in a year. For middle school, I assume that any teacher matched to more than 200 students in one year is a misclassification and code these individuals as not being matched to any teacher.

Elementary-To-Middle School Transitions Figure H2a illustrates how the make-up of elementary schools in New York City changed over time. In the early 1990s, the majority of elementary schools were K-6, with only about one-third being K-5 and very few being K-8. Over time, this changed significantly as the fraction of K-6 schools decreased precipitously while both the fraction of K-5 and K-8 schools increased. I therefore focus on the elementary-to-middle school transition by running the regressions using all students in the first year at their middle school, rather than focusing on a particular grade. In practice, this means that not all students in the sample are sixth graders (because some middle schools consist of grades 7-8) and not all sixth graders are in the sample (because some students attend either K-8 or K-6 schools). Figure H2c shows that almost no non-sixth graders are included by the end of the period, when K-6 elementary schools and 7-8 middle schools pretty much vanished in New York City. Even then, however, not all sixth grade students are included, primarily because of the increase in K-8 schools.

Estimating Teacher Value-Added The Chetty, Friedman and Rockoff (2014*a*) method for estimating VA proceeds in four main steps. The first is to remove determinants of student i's test score that a teacher cannot affect. This is done by regressing student i's year t test score, denoted as $y_{i,t}$, on a vector of student i observables, denoted as $X_{i,t}$. Importantly, $X_{i,t}$ contains cubic functions of student i's lagged test scores. In my data, adding additional controls do little to change the VA estimates, a finding the resembles that of Chetty, Friedman and Rockoff (2014*a*).[56] The regression to estimate the effect of $X_{i,t}$ on $y_{i,t}$ includes teacher fixed effects, which removes the possibility that the estimate is biased by teachers sorting based on the X's. Once β is estimated, I construct student level residuals $y^*_{i,t} = y_{i,t} - \hat{\beta}X_{i,t}$.

Once these student-level residuals are constructed, the next step is to aggregate them to the teacher-year level. For teacher j, I denote her year t measure as $A_{j,t}$.

[56]My main specification includes only cubic functions of a student's lagged math and English test scores, but both the magnitude of my coefficients and their t-statistics increase slightly when including all the controls used in Chetty, Friedman and Rockoff (2014*a*).

To be clear, $A_{j,t}$ is just the sum of her students' residuals: $A_{j,t} \equiv \sum_{\forall i \in c(j,t)} y_{i,t} - \hat{\beta} X_{i,t}$, where $c(j,t)$ indicates the set of students that teacher j teaches in year t.

The two steps above provide me with a measure $A_{j,t}$ for every teacher-year. This measure combines teacher j's effect on her year t students with all the other uncontrolled for determinants of her student's test score residuals. To remove the contemporaneous error terms from $A_{j,t}$, the Chetty, Friedman and Rockoff (2014*a*) estimation technique uses the inter-temporal correlation between $A_{j,t}$ and $A_{j,-t}$, where $A_{j,-t}$ is a vector of every $A_{j,t'}$ measure such that $t' \neq t$. In particular, it assumes a stationary process for both the true teacher VA and for the student-level error terms and estimate $Cov(A_{j,t}, A_{j,t-s}) \equiv \sigma_{As}$ for all $s \in \{1,2,3,4,5,6,7\}$, assuming that the correlations stabilize after seven years.

Once these inter-temporal covariances are estimated, the last step is to predict teacher j's value of $A_{j,t}$ using $A_{j,-t}$. This is done using the estimates of $\hat{\sigma}_{As}$ and the measures in $A_{j,-t}$. These predictions become the estimated VA of teacher j in year t, which I will denote as $\hat{\mu}_{j,t}$. As an example, suppose that teacher j was teaching in New York City from 2005 to 2009. Then teacher j's estimated VA in 2007 is:[57]

$$\hat{\mu}_{j,2007} = \hat{\sigma}_{A2} A_{j,2005} + \hat{\sigma}_{A1} A_{j,2006} + \hat{\sigma}_{A1} A_{j,2008} + \hat{\sigma}_{A2} A_{j,2009} + \hat{\sigma}_{A3} A_{j,2010} \tag{A1}$$

and teacher j's estimated value-added in 2008 is:

$$\hat{\mu}_{j,2008} = \hat{\sigma}_{A3} A_{j,2005} + \hat{\sigma}_{A2} A_{j,2006} + \hat{\sigma}_{A1} A_{j,2007} + \hat{\sigma}_{A1} A_{j,2009} + \hat{\sigma}_{A2} A_{j,2010} \tag{A2}$$

Although this ensures that the test scores of teacher j's year t students have absolutely no impact on $\hat{\mu}_{j,t}$, since I am interested in cohort VA changes, that is not quite enough. To see why, notice that:

$$\begin{aligned}\hat{\mu}_{j,2008} - \hat{\mu}_{j,2007} =& (\hat{\sigma}_{A3} - \hat{\sigma}_{A2}) A_{j,2005} + (\hat{\sigma}_{A2} - \hat{\sigma}_{A1}) A_{j,2006} \\ &+ \hat{\sigma}_{A1} A_{j,2007} - \hat{\sigma}_{A1} A_{j,2008} \\ &+ (\hat{\sigma}_{A1} - \hat{\sigma}_{A2}) A_{j,2009} + (\hat{\sigma}_{A2} - \hat{\sigma}_{A3}) A_{j,2010}\end{aligned}$$

The most concerning aspect is the middle term, $\hat{\sigma}_{A1} A_{j,2007} - \hat{\sigma}_{A1} A_{j,2008}$, which implies that the change in a teacher's own VA is inversely correlated with changes in the quality of her students. When estimating the VA components of my main measure, I therefore follow Chetty, Friedman and Rockoff (2014*a*) by excluding two years of students. In the example above, that would mean estimating $\mu_{j,2008}$ by excluding both the students teacher j had in 2008 and in 2007, and estimating $\mu_{j,2007}$ by excluding the students teacher j had in 2007 and in 2008. As is clear from the above expression, excluding both 2008 and in 2007 alone is not enough to

[57]For a more detailed discussion of how to account for the fact that teachers teach for a different number of years and for the fact that the variance of $A_{j,t}$ differs for every (j,t) pair, see Appendix A of Chetty, Friedman and Rockoff (2014*a*).

ensure $\mu_{j,2008} = \mu_{j,2007}$, since recent years are weighted more strongly than distant years when estimating $\mu_{j,t}$. I show in Section C that averaging the σ-weights to make $\mu_{j,2008} = \mu_{j,2007}$ does not affect the results.

Imputing Missing Teacher Value-Added The downside of the above method for estimating teacher VA is that it means that I cannot estimate VA for teachers who are in the data for a short period of time. For example, in the example above it is impossible to estimate teacher j's VA while omitting her students in 2007 and in 2008 if those were the only two years she taught in New York City. In addition, the data is missing teacher VA estimates for a large fraction of students in the early-to-mid-1990s, as shown in Figure H1. This is due mostly to the fact that the data system used to keep track of student to teacher matches was slowly phased in during this time, but means there are additional students for whom I cannot measure their teacher's VA. This complicates the analysis because calculating how the average teacher VA in the highest grade at elementary school e changed between year $t-1$ and year $t-2$ requires VA estimates for all teachers. This means I need to impute VA for unmatched students and teachers who are only in the data during years $t-1$ and $t-2$. In the main specification, I do so by assuming that the teacher VA at elementary school e for unmatched students and teachers for whom I'm unable to estimate teacher VA is the same in year $t-1$ as in $t-2$; one reason this would be true is if I used the common approach of imputing missing VA as being the sample mean, and it can also hold under other imputation approaches as well. Here I show that the results are robust to other imputation approaches.

The first alternative approach is to assume that the change in teacher VA at elementary school e for unmatched students and teachers for whom I'm unable to estimate teacher VA is the same as the change in teacher VA at the elementary school for teachers with VA. That is, instead of assuming that the missing VA does not change between $t-1$ and $t-2$, it assumes that the missing VA changes identically to the non-missing VA. Table H1a shows that this imputation approach gives similar coefficients to the main imputation approach.

Another is to not omit any years when estimating teacher VA, i.e. in the example above including students from both 2007 and 2008 when estimating the teacher's value-added for either year. Although this potentially biases the VA estimates themselves, it also reduces the need for imputation. Table H1b shows that the results using this approach are again similar to the baseline results shown in Table 2.

As a final robustness check, I restrict the sample to post-1998 data, when the number of missing VA estimates is much lower than in the pre-1998 data. The results of the baseline specification using post-1998 data are shown in Table H1c. Again, these results closely match Table 2, which I consider to be further evidence that the imputation approach is not driving the results.

B. Leave-Out Mean Vs. Shrunken Leave-Out Mean

As discussed in Section II, I create a measure that shrinks the leave-out mean toward zero in proportion to the size of the left-out group. In this section, I show theoretically this approach is conceptually similar to using a traditional leave-out mean and gives rise to the same point estimates when the size of the left-out subgroup is fixed. I also show that the approach used in this paper provides more precise estimates when the size of the left-out subgroup varies. Although this motivates the approach I use in the paper, I also show empirically that using a traditional leave-out mean gives me similar results in the context of this paper.

B1. Theoretical Results

To set up the theoretical results, I start by simplifying the notation a bit. I'll denote the outcome we care about for some subgroup i of group j as y_{ij}. Likewise, I'll denote the average covariate for that group as X_{ij} and the average of the other subgroups as $X_{(-i)j}$. In addition, the fraction of the group j that is part of subgroup i is denoted $\alpha_{ij} \in (0, 1)$. Finally, I will denote the overall average of X for group j as X_j. Thus, $X_j = \alpha_{ij} X_{ij} + (1 - \alpha_{ij}) X_{(-i)j}$.

In what follows, I will assume that an individual's value of y is affected by her peers' average value of X, plus an error term, so that:

$$y_{ij} = \alpha + \beta X_j + \epsilon_{ij} \tag{B1}$$

If X_{ij} has a direct effect on y_{ij}, the β parameter in Equation (B1) will not identify the desired peer effect estimate since X_j will be correlated with ϵ_{ij}. A natural fix for this is to only use variation of X_j that comes from $X_{(-i)j}$ to identify β. This can be done in three ways; in what follows I show that all three approaches converge to β. The first and most common approach is to use a conventional leave-out mean, i.e. $X_{(-i)j}$, as an instrument for X_j. The second approach is to use a shrunken version of the leave-out mean, i.e. $(1 - \alpha_{ij}) X_{(-i)j}$, as an instrument for X_j. The third approach is to use the shrunken leave-out mean in an OLS regression, rather than as an instrument for X_j. This last approach is the one I use in this paper.

Since the question here is not whether it is possible to identify peer effect, but instead how best to identify the peer effect, I will assume that the variation is sufficient to estimate the peer effect. This involves three main assumptions: first, that $X_{(-i)j}$ is uncorrelated with X_{ij}; second, that $Var(X_{(-i)j}) > 0$; third, that $Cov(X_{(-i)j}, \epsilon_{ij})$. The first assumption is mainly one of convenience, and allows the analysis below to consider univariate regressions instead of multivariate regressions in which both $X_{(-i)j}$ and X_{ij} are included as covariates. The second assumes that there is meaningful variation in $X_{(-i)j}$, and the third that this variation is exogenous.

I will also assume that the size of the left-out group (i.e. α_{ij}) is uncorrelated with the other variables. If this assumption fails, the choice of whether to use the shrunken or unshruken leave-out mean would be determined by which measure is truly exogenous, which would depend on the empirical context. I focus here on a case when the assumption does hold and show in the next subsection that the three approaches all give similar estimates in the context of this paper. It is also worth noting that this assumption is trivially true when the group sizes are all identical, such as when the subgroup i consists of a single individual. Thus, a corollary of the results below is that the three approaches give rise to the same coefficient estimates when the group sizes are all identical.

Given these assumptions, it is straightforward to show that the three approaches converge to the same parameter estimates. Because the regressions are univariate, the three parameter estimates are as follows:

$$\hat{\beta}_1 = \frac{Cov\Big(y_{ij}, X_{(-i)j}\Big)}{Cov\Big(X_j, X_{(-i)j}\Big)} \tag{B2}$$

$$\hat{\beta}_2 = \frac{Cov\Big(y_{ij}, (1-\alpha_{ij})X_{(-i)j}\Big)}{Cov\Big(X_j, (1-\alpha_{ij})X_{(-i)j}\Big)} \tag{B3}$$

$$\hat{\beta}_3 = \frac{Cov\Big(y_{ij}, (1-\alpha_{ij})X_{(-i)j}\Big)}{Var\Big((1-\alpha_{ij})X_{(-i)j}\Big)} \tag{B4}$$

Plugging in Equation (B1) for y_{ij} and the fact that X_{ij} and $X_{(-i)j}$ are uncorrelated, some more algebra shows that:

$$\hat{\beta}_1 = \beta + \frac{Cov\Big(\epsilon_{ij}, X_{(-i)j}\Big)}{Cov\Big(X_j, X_{(-i)j}\Big)} \tag{B5}$$

$$\hat{\beta}_2 = \beta + \frac{Cov\Big(\epsilon_{ij}, (1-\alpha_{ij})X_{(-i)j}\Big)}{Cov\Big(X_j, (1-\alpha_{ij})X_{(-i)j}\Big)} \tag{B6}$$

$$\hat{\beta}_3 = \beta + \frac{Cov\Big(\epsilon_{ij}, (1-\alpha_{ij})X_{(-i)j}\Big)}{Var\Big((1-\alpha_{ij})X_{(-i)j}\Big)} \tag{B7}$$

The assumptions that $Cov\Big(\epsilon_{ij}, X_{(-i)j}\Big) = 0$ and that α_{ij} is uncorrelated with the other variables are enough to show that all three estimates converge to β.

Still, the question remains, which has the smallest asymptotic variance? To answer this, I will make an additional assumption that ϵ_{ij} is homoscedastic.[58]

[58]Since y_{ij} is an average of the values of y for all individuals in subgroup i, this seems to be a

Given the assumption of homoscedastic error terms and the assumptions above, some algebra shows that:

$$\begin{aligned} Var(\hat{\beta}_1) =& Cov\Big(X_j, X_{(-i)j}\Big)^{-1} \cdot Var\Big(X_{(-i)j}\Big) \cdot Cov\Big(X_j, X_{(-i)j}\Big)^{-1} \\ =& \Bigg[\Big(1 - \mathbb{E}[\alpha_{ij}]\Big)^2 \cdot Var\Big(X_{(-i)j}\Big)\Bigg]^{-1} \end{aligned}$$

and

$$\begin{aligned} Var(\hat{\beta}_2) = Var(\hat{\beta}_3) =& \Bigg[Var\Big((1-\alpha_{ij})X_{(-i)j}\Big)\Bigg]^{-1} \\ =& \Bigg[\Big(1 - \mathbb{E}[\alpha_{ij}]\Big)^2 \cdot Var\Big(X_{(-i)j}\Big) + Var\Big(1-\alpha_{ij}\Big)Var\Big(X_{(-i)j}\Big)\Bigg]^{-1} \end{aligned}$$

Thus, as long as there is some variation in the group sizes, i.e. that $Var\Big(1-\alpha_{ij}\Big) \neq 0$, using the shrunken leave-out mean either as an OLS or IV gives rise to more precise estimates. This is because accounting for variation in the group sizes means that there is a closer correspondence between the variable of interest and the instrument, which leads to a stronger first-stage and therefore more precise estimates.

B2. Leave-Out Mean Estimates

The previous subsection used theory to motivate the use of the shrunken leave-out mean. In this subsection, I show that I get similar estimates regardless of whether I use the shrunken leave-out mean or the conventional leave-out mean.

The first two columns of Table H2a report the results from IV regressions, where either the shrunken leave-out mean or conventional leave-out mean are used as instruments for the overall mean. Using the notation of the paper, the shrunken leave-out mean is $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ and the conventional leave-out mean is $\frac{\Delta\mu^{peer}_{c(e,m,t),s,t-1}}{1-\alpha_{e,m,t}}$. As is expected from the above results, these two estimates are quite similar. The next two columns report the results from OLS regressions, using the weighted and conventional leave-out mean. Again, the third column is quite similar to the first two. However, the OLS regression using the conventional leave-out mean is smaller. It is straightforward to show that, under the assumptions

problematic assumption as it is likely that the error component will depend on the number of individuals in subgroup i. However, there is often a component of the error term that affects all members of subgroup i. This assumption then is roughly that this common shock to the subgroup i is both homoscedastic and large enough to subsume the idiosyncratic components. Even if that is not the case, the intuition presented here should still hold; however, to minimize the asymptotic variance the larger subgroups should be given heavier weight.

in Section B.B1, this coefficient should be equal to the coefficients in the other three columns times $1-\mathbb{E}[\alpha_{e,m,t}]$. In the elementary-to-middle school sample, $1-\mathbb{E}[\alpha_{e,m,t}] \approx 0.8$ which is approximately the ratio between the estimated coefficient in Column (4) and those in Columns (1) - (3).

In Table H2a, the standard errors in Columns (1) - (3) are similar. This is largely because there is a relatively small amount of variation in $1-\alpha_{e,m,t}$ when focusing on the elementary-to-middle school transitions. Table H2b reports the same four coefficients when using *all* school transitions to estimate the spillover effect. Again, the point estimates of the first three columns are similar. However, in this case, there is enough variation in $1-\alpha_{e,m,t}$ that using the shrunken leave-out mean either as an instrument or via OLS gives much more precise estimates. In addition, because $1-\mathbb{E}[\alpha_{e,m,t}]$ is much smaller in this case, the difference between the point estimates in the fourth column and the other three is much larger.

C. Robustness Checks

C1. Other VA Measures

Given the high correlation between different value-added measures, it is unlikely that the results would be affected by the VA model. This subsection ensures that is the case by running the same baseline regression, using different VA models. First, I use VA that are estimated the same control vector as in Chetty, Friedman and Rockoff (2014*a*). In addition to a student's lagged test scores, this specification includes student-level information on their: gender, lagged days absent, relative age, race, absences, and discipline incidents. It also includes information on whether the student has repeated the grade, whether or not he or she is classified as an English Language Learner, and whether or not he or she is classified as having a learning disability. This control vector also includes interactions of the cubic function of a student's test scores with the student's grade, to allow test score growth to differ depending on the student's age. It includes classroom-level averages of all the previous controls as well as controls for the number of other students in the class. The results from this specification are reported in Table H3a. As can be seen, they closely match the results in Table 2.

As discussed in Section I.B most, but not all, of the variation in teacher VA is across-teachers rather than within-teachers. The fact that there is some drift in the VA estimates means that not all of the variation used to identify the spillovers comes from teacher transitions. Given the low amount of drift, it is unlikely that affects the results, and Table H3b verifies that this the case. It runs specifications identical to Equation (2), but averages the teacher VA estimates to ensure all of the variation in $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ comes from teacher transitions rather than within-teacher drift. Again, the results in Table H3b are similar to those in Table 2.

C2. IV Regressions

As discussed in Section II, the measure I use in each regression excludes variation in the previous teacher quality of one's peers that is generated by changes in the way teachers get assigned to students within the elementary schools or the way those students choose middle schools. Excluding this variation increases the likelihood that the variation used to identify the spillovers is exogenous, but it also suggests that an IV approach may be appropriate. In this, I use the main measure as an instrument for changes in the actual previous teacher quality of a cohort's peers, instead of it directly in an OLS regression. To conduct the IV approach, I first average the previous teacher value-added over all the new students at a middle school in a given year, regardless of the elementary school they attended. I then calculate the actual change in previous teacher quality of a cohort's peers as the year-to-year difference in this measure. Finally, I run a 2SLS regression, using the main measure $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ as an instrument for the actual change.

As shown in Table H4, the coefficients from this procedure gives similar coefficients as when the main measure is used directly as in Equation (2). This is because the major difference between $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ and the actual change in the previous teacher quality of a cohort's peers is the fact that $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ excludes individuals who attended elementary school e, and that is already accounted for in the construction of $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$. See Section II and B for more discussion of this fact. Thus, the only remaining differences between $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ and the actual change in the previous teacher quality of a cohort's peers are driven by changes in the way teachers are sorted to students at the neighboring elementary schools and how students from the neighboring elementary schools sort into middle schools, and these patterns stay quite constant over time.

C3. Pseudo-Zoned Schools

As discussed in Section II, the measure I use in each regression is not affected by changes in the way individuals at a student's neighboring elementary schools get sorted to middle schools. Yet it is affected by where the student herself attends middle school. Although most students attend their closest middle school, students do have the flexibility to choose their middle school in the later years of the analysis. Since I always compare how students score relative to those who attended the same elementary and middle school in the previous year, it is unclear how, or if, this choice would bias the results. Regardless, this section ensures that students' choice of middle school does not have any effect on the results presented, nor does any grade repetition or endogenous parental movements.

In theory, the best way to handle this choice is to construct the measure that assumes that all students attend their zoned middle school, i.e. the school that they are defaulted in to. Since I do not have this information I instead use a different approach that has a similar flavor as using a student's zoned school, which involves using what I call a student's "pseudo-zoned school." This technique involves constructing the main covariate under the assumption that each student has the same probability of attending each middle school as the average person at her elementary school. This ensures that the main covariate of interest is not affected by the student's choice of middle school to attend.

More specifically, this means the measure, denoted as $\Delta\mu^{pseudo\_zoned}_{e,t-1}$, becomes:

$$\Delta\mu^{pseudo\_zoned}_{e,t-1} = \sum_{\forall m} \beta_{e,m,t} \sum_{\forall e' \neq e} \alpha_{e',m,t} \Delta\mu_{e',t-1} \tag{C1}$$

where $\Delta\mu_{e',t-1}$ is change in the average teacher VA in time $t-1$ for the teachers at the highest grade of elementary school e', $\alpha_{e',s'_m,t}$ is the fraction of students at middle school m that attended elementary school e', and $\beta_{e,m,t}$ is the fraction of students who attended elementary school e that move on to attend middle school m. Note that $\Delta\mu^{pseudo\_zoned}_{e,t-1}$ is only a function of the elementary school the student went to, and not a function of the middle school the student attended, which explains the different subscripts. This ensures that the middle school the student attended has no effect on the measure. I then run the same specification outlined in Equation (2), but now use $\Delta\mu^{pseudo\_zoned}_{e,t-1}$ as an instrument for $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$.

The regression results are demonstrated in Table H5. The point estimates are a bit larger than those in Table 2, but in the same general range. The use of an instrument also increases the standard errors.

C4. Within-Group Spillovers

In Section V, I use the fact that the flow rates from elementary schools to middle schools differ slightly by subgroups and the fact that within each elementary school the students had slightly different teachers to show that the spillovers occur within-subgroups, where the subgroups are defined by a student's race and gender. Here, I show that the same result holds when ignoring variation that comes from the fact that within each elementary school the students have slightly different teachers. I also show that the same result holds when defining subgroups based on whether or not the student is classified as being an English Language Learner (ELL).

The regressions I run are identical to the one described in Section V, with the only difference being that the covariates are calculated slightly differently. In the main section of the paper, the measures were constructed using the same form as Equation (1), where both the feed rates (i.e. $\alpha_{e',m,t}$) and the elementary school teacher value-added changes (i.e. $\Delta\mu_{e',t-1}$) were allowed to vary based on the subgroup considered. When defining the subgroup based on race, I show in Table H6a that I find similar results as before when the measures are constructed using only variation in the feed rates to identify the difference spillovers.

Allowing the α's to vary is not enough to separately determine whether the spillovers occur within- or across-genders. Because of this, I add an additional source of variation: the fact that some teachers might be better at teaching male students than female students, or vice versa.[59,60] Unsurprisingly, teachers who are good at increasing VA for one subgroup of students tend to be good at increasing VA for others, a finding illustrated for the subgroups of gender and ELL status in Figures H3a and H3b, respectively. Yet there are persistent differences in how effective teachers are for different subgroups. This means that the $\Delta\mu_{e',t-1}$ will differ across subgroups, even ignoring the fact that the members of the different subgroups may have had different teachers at elementary school e'. Tables H6b and H6c use variation in the feed rates and in the teachers' differential ability at teaching members of the subgroup (but not variation in which teachers members of the subgroup had) to separately estimate the within-group versus across-group spillovers. Although I am not able to reject the fact that these spillovers are different when defining groups based on whether a student in classified as an English Language Learner, the point estimates suggest that the spillovers are much larger within-groups. In contrast, when defining subgroups based on gender, the two point estimates appear similar across all of the specifications.

Overall, I view the results of Table H6a, H6b, and H6c as providing some support for the robustness of the findings reported in Tables 10 and 11 that students are only affected by the quality of the teachers who previously taught the other students at the school who are similar to themselves.

C5. Dynamic Analysis

So far, I've focused on how teachers affect the subsequent classmates of their students by focusing on the year after the teacher taught the student. It's quite possible, however, that the teacher affects the classmates of their students two years after. In this section, I test whether this is the case. Table H7 presents regressions similar to the ones reported in Table 2; however, I now include both the teacher VA of a student's peers' prior teachers and the teacher VA of a student's peers' teachers two years prior. Not surprisingly, the effect of a student's peers' prior teachers on her test scores is larger than the effect of a student's peers' two years prior teachers, but the student's peers' two years prior teachers do appear to have an effect on her test scores.

[59]Most of the evidence on whether a teacher's effectiveness differs across genders has focused on gender bias. These biases have been shown to have a lasting negative effect in Lavy and Sand (2015). Evidence that some teachers' effectiveness is different for ELL students than for non-ELL students is shown in Loeb, Soland and Fox (2014).

[60]I do not replicate this analysis for race, because it is generally impossible to estimate each teacher's VA for every race because of limited sample size.

D. Using the Entire Sample of Individuals

D1. Estimates Using the Full Sample of Students

In the main section of the paper, I focus exclusively on the elementary-to-middle school transition of students, ignoring the fact that some students transfer between schools in every grade. In this section of the appendix, I show that the main results are robust to using the full sample of individuals. For example, Table H8 is identical to Table 2, with the exception being that it now reports the coefficients that arise from estimating Equation (2) using all grades instead of focusing on the elementary-to-middle school transition. The point estimates in Table H8 are a lower than those in Table 2, but are still statistically significant and large enough to be economically meaningful. The same is true when estimating the spillover effects separately by subject as shown in Table H9.

Similarly, Tables H10, H11, and H12 show the falsification and placebo tests when using the full sample. Although including the non-elementary-to-middle school transitions adds additional endogeneity concerns, these concerns generally do not appear to be a major issue given that nearly all of the coefficients in the falsification and placebo tests are not statistically significant.

Finally, Tables H13, H14, H15, and H16 show that all the conclusions from Section V are replicated when using the full sample. In addition, the increased sample size increases the statistical power, which means that I reject more of the tests that the estimated coefficients are equal.

D2. Specification Tests

As well as adding additional statistical power to the estimates, there is an additional benefit to using the entire sample of individuals in the regressions. Unlike in the case when I focus on the transition from elementary-to-middle school, there is now extensive variation in the fraction of individuals who previously attended the same school as an individual, i.e. in the $\alpha_{e,m,t}$ terms. This presents the opportunity to conduct two specification tests, as well as regressions that combine the first specification tests with the falsification and placebo tests discussed in the paper.

First Specification Test

If the correlation I have demonstrated so far is causal, changes in the average teacher VA at students' neighboring elementary schools will affect them more when more of their middle school peers come from the neighboring elementary schools. This section provides more evidence in support of the identification assumption by testing this proposition directly.

To do so, first recall that the main measure I use is defined as:

$$\Delta\mu^{peer}_{c(e,m,t),s,t-1} \equiv \sum_{\forall e' \neq e} \alpha_{e',m,t} \cdot \Delta\mu_{s,e',t-1} \tag{D1}$$

where $\Delta\mu_{e',t-1}$ measures how the average teacher value-added at elementary school e' changed between year $t-1$ and year $t-2$, and $\alpha_{e',m,t}$ is the fraction of students at middle school m who came from elementary school e'. As discussed in Section B, an alternative is to a construct a more traditional leave-one-out mean by dividing $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ by $(1-\alpha_{e,m,t})$, which corresponds to the fraction of students at middle school m who did not attend elementary school e. If there is enough variation in $(1-\alpha_{e,m,t})$, it is possible to run the following regression:

(D2)

$$\Delta Y_{c(e,m,t)s,t} = \alpha+\gamma_0\cdot\left(\frac{\Delta\mu^{peer}_{c(e,m,t),s,t-1}}{1-\alpha_{e,m,t}}\right)+\gamma_1\cdot(1-\alpha_{e,m,t})+\gamma_2\cdot\Delta\mu^{peer}_{c(e,m,t),s,t-1}+\Delta\epsilon_{c(e,m,t)s,t}$$

This specification separately controls for average teacher VA at the other elementary schools that feed the cohort's middle school, the fraction of students in the cohort's middle school who did not attend the cohort's elementary school, *and* for the interaction between these terms.[61] If the earlier results are indeed due to the changes in a student's peers' underlying ability caused by the quality of her teacher, the interaction term should matter, while the un-interacted leave-out average term will not.

Estimating Equation (D2) requires there to be sufficient variation in $(1-\alpha_{e,m,t})$. Since most students in New York City attend middle schools with students who previously attended many different elementary schools, the value of $(1-\alpha_{e,m,t})$ is close to one for nearly every cohort when restricting the sample to the elementary-to-middle school transitions that I use in the main body of the paper. This means that I cannot separately estimate $\frac{\Delta\mu^{peer}_{c(e,m,t),s,t-1}}{1-\alpha_{e,m,t}}$ and $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ when restricting the sample, as the correlation between the two variables is over 0.98. When using the full sample of individuals, however, this correlation drops to 0.60 and it becomes possible to estimate Equation (D2). It is also possible to estimate a more flexible specification which relaxes the assumption that the spillover effects scale linearly with the size of the left-out cohort. This specification is defined as:

$$\text{(D3)} \qquad \Delta Y_{c(e,m,t)s,t} = \sum_{k=1}^{10}\alpha_k I_k + \gamma_k I_k\frac{\Delta\mu^{peer}_{c(e,m,t),s,t-1}}{1-\alpha_{e,m,t}} + \Delta\epsilon_{e,m,t}$$

where I_k is an indicator variable that equals one if and only if $(1-\alpha_{e,m,t}) \in \left[\frac{k-1}{10}, \frac{k}{10}\right]$.

Table H17 shows the estimated coefficients from Equation (D2) and Figure

[61] As mentioned earlier, by necessity both the average teacher VA and the α's exclude individuals who did not previously attend a public school in New York City. This is a relatively small fraction of the students, however, and so will not affect the results unless average teacher VA changes at these schools is strongly correlated with average teacher VA changes of the other elementary schools that feed the middle school. Given the fact that the elementary school VA changes are uncorrelated within New York City, as shown in Table 2a, this seems unlikely.

H4 plots the γ_k coefficients from Equation (D3), along with the distribution of $(1-\alpha_{e,m,t})$ and the estimated effect when assuming that the effect scales linearly. Both specifications demonstrate that as the percentage of student's peers who are affected by changes in teacher VA increases, the effect on the student increases. In Table H17 the interaction term is positive, statistically significant, and similar to the results in Table 2. The coefficient on the unweighted change, in contrast, is a quite precisely estimated zero. This provides one more piece of evidence that the results are indeed due to the proposed causal mechanism.

Second Specification Test

Another testable implication of the spillovers is that the effect of a cohort's *own* previous teachers' VA should be larger if more of their middle school peers came from the same elementary school. This is because the effect of cohort's own previous teachers' VA combines both the fade-out of the direct effect and the presence of the spillover effects. To test this, I run the following regression:

$$\Delta Y_{c(e,m,t),s,t} = \alpha + \beta_0 \cdot \Delta\mu_{e,t-1} + \beta_1 \cdot \Big[\alpha_{e,m,t}\Delta\mu_{e,t-1}\Big] + \Delta\epsilon_{e,m,t} \tag{D4}$$

Like before, $\Delta\mu_{e,t-1}$ measures the previous teacher VA for cohort $c(e,m,t)$. Here β_1 tests whether this scales with the size of the cohort, measured by $\alpha_{e,m,t}$. The results, shown in Table H18, suggest that the effect does scale with the size of the cohort. Note that this is in contrast to estimates on the effect of a cohort's current teachers' VA, which does not depend on the size of the cohort or the level of analysis. This adds to the evidence that the spillovers discussed in this paper exist.

Combining the Specification Tests and the Placebo Tests Another possibility is to combine the specification test and the placebo tests, although this again is only possible when using the full sample of individuals. This clearly illustrates whether or not there is any systematic correlation between the leave-out average of a peers' previous teacher's VA, the fraction of peers who previously attended the same school, their interaction, and the variables that should be exogenous. To do so, I first replicate the specification described by Equation (D3), but use demographic variables instead of test scores as the outcome measure. The results of these regressions are shown in Table H19. Of the 18 coefficients that were estimated, only one was statistically significant.

I next combine the first placebo test with the specification test. For this test, I again replicate the specification described by Equation (D3), but now include both lag and lead values of each covariate. As shown in Table H20, the only variable that is consistently positive and statistically significant is the current measure of the interaction between the leave-out average of a peers' previous teacher VA and the fraction of the peers that are affected by this measure, i.e. the main variable I use in all my specifications $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$. I consider this to be more evidence that the identification approach is valid and the main specification is correct. Finally, I combine the second placebo test with the specification test, which is shown in Table H21. Again, all of the point estimates are quite small and all but one of the coefficients are statistically insignificant.

E. Placebo Test Details

As discussed in Section III.B, one of the placebo tests is designed to estimate the correlation between a student's test scores and her *future* peers' previous teachers' VA using a similar procedure as my main specification. If no students switched schools other than the elementary-to-middle school transition, this specification would be nearly identical to Equation (2). The only difference would be that now $\Delta\mu^{peer}_{c(e,m,t),s,t-1}$ would measure changes in the teacher value-added in the next-to-last grade at the neighboring elementary schools and $\Delta Y_{c(e,m,t),s,t}$ would measure changes in the student test scores in the last grade at their elementary school. In practice, the number of students who change schools between other grades is small relative to those who switch between elementary and middle school, but some do switch. This section discusses how to account for these.

The first step is to estimate $\alpha'_{e,m,t}$, which is the fraction of students at middle school m who attended elementary school e two years before the transition. This differs from the measure $\alpha_{e,m,t}$ used in the rest of the paper only in that it measures where the student attended elementary school two years before moving to middle school, rather than the year before. Not surprisingly the two measures are strongly correlated, with the estimated correlation being over 0.95.

Given these weights, $\alpha'_{e,m,t}$, I then construct a measure for the lagged teacher VA of cohort $c(e,m,t)$'s future peers, denoted as $\mu^{placebo}_{c(e,m,t),s,t-2}$, similar to before:

$$\Delta\mu^{placebo}_{c(e,m,t),s,t-2} = \sum_{\forall e' \neq e} \alpha'_{e,m,t} \Delta\mu_{e',t-2} \tag{E1}$$

Again, the only difference is that $\Delta\mu_{e',t-2}$ is now measured as the change in the average teacher value-added in the penultimate grade of elementary school e', rather than the final grade of elementary school e'. Note also that the term is indexed by $t-2$ since it measures the teacher quality in the neighboring elementary schools two years before cohort $c(e,m,t)$ makes the transition from elementary school to middle school.

Two additional adjustments are needed to account for the fact that some students switched schools in years other than the elementary-to-middle school transition. First, I now define a cohort as a group of students who attended the same school three years in a row (elementary school the first two years and middle school the third year). Second, the switching of schools means that some individuals who went to the same middle school as a student and did not go to school with him or her two years prior, did go to school with him or her in the last year of elementary school. These students appear in the placebo measure, causing some positive correlation between it and a student's test scores. To account for this, I include an additional control in the placebo tests: the true lagged teacher VA of a student's peers. This makes very little change to the results, but makes the coefficients a more accurate test for the existence of spurious correlation.

Given these changes in the definition of cohorts, the placebo regression stays nearly the same as Equation (2):

(E2)

$$\Delta Y_{c(e,m,t),s,t-1} = \alpha + \gamma_{placebo}\Delta\mu^{placebo}_{c(e,m,t),s,t-2} + \gamma\Delta\mu^{peer}_{c(e,m,t),s,t-2} + \Delta\epsilon_{c(e,m,t),s,t-1}$$

The time index on all of these variables shift back one year, since they measure outcomes in the year before cohort $c(e, m, t)$ makes the transition from elementary school to middle school.

Another placebo test mentioned in Section III.B is designed to estimate the correlation between a student's test scores and her future peers' current teachers' VA. This is done identically to the method discussed above with one exception. Instead of using the student's previous teachers to construct the placebo measure, I use the student's current teachers to do so.

F. Direct Value-Added Fade-Out

As discussed in Section IV, to compare the direct value of a teacher to the indirect value, it is important to know the effect of a teacher on her students in the year after the teacher taught them. This section calculates this fade-out, as well as confirms existing evidence that an increase in a teacher's value-added increases her students' contemporary test scores one-for one.

To do so, I aggregate both the test scores and VA to the cohort-level and then run a regression similar to Equation (2). This estimates the effect of changes in teacher VA, either contemporary or previous, on changes in test scores.

The results are presented in Table H22. The first column replicates the results of Chetty, Friedman and Rockoff (2014*a*), showing that the direct effect of a teacher on their own students is equal to the teacher's VA. The second column shows how this effect fades out after one year. It shows that an increase in the teacher's VA increases the students test scores by 0.55 in the subsequent year, which corresponds to the results presented in Figure 4 of Chetty, Friedman and Rockoff (2014*b*).

G. Measurement Error Simulation

Feld and Zölitz (2017) show that when peer groups are not formed under random assignment, the presence of measurement error in the covariates can increase the coefficient estimates rather than attenuate them. As shown in Figure 2a, changes in teacher VA at a cohort's neighboring schools is more or less uncorrelated with changes in teacher VA at a cohort's own elementary school, which suggests that this complication in measuring peer effects is not an issue in this context. Like Feld and Zölitz (2017) suggest, however, I ensure that it is not an issue by conducting a Monte Carlo simulation.

To conduct the simulation, I start with a dataset containing every teachers' estimated VA. I then add a randomly drawn error term to each teachers' VA. The error terms are drawn from a normal distribution with a mean of zero and a standard deviation that I vary depending on the specification. I then merge these new teacher VA estimates, which now contain more measurement error than the original VA estimates, back to the student data.

At this point, I proceed in the same way as described in Section II. This involves first aggregating the individual data to the cohort-level, while constructing both a cohort's own average previous teacher VA and a cohort's peers' average previous teacher VA. I then estimate the effect of a cohort's peers' average previous teacher VA by running the regression defined in Equation (2). For the Monte Carlo simulation, I run the specification that also controls for the cohort's own average previous teacher VA so that, when there is no added measurement error, the approach would give the coefficient reported in Column (2) of Table 2.

I conduct the above process using thirteen different values for the standard deviation of the error term, ranging from 0 to 0.3. For each value of the standard deviation, I conduct five simulations to ensure that idiosyncrasies in the draw of the error terms do not affect the results. The average of these five simulations is reported in Figure H5. As can be seen, in this context more measurement error in the teacher VA estimates causes attenuation in the estimated peer effect.

H. Appendix Figures and Tables

H1. Appendix Tables

Table H1—Other Imputation Approaches

(a) Assume Missing VA Changes Are Identical to Non-Missing VA Changes

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.431	0.404	0.340	0.504
	(0.131)	(0.130)	(0.112)	(0.138)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Imputation Approach	Missing VA Changes Equals Non-Missing VA Change	Missing VA Changes Equals Non-Missing VA Change	Missing VA Changes Equals Non-Missing VA Change	Missing VA Changes Equals Non-Missing VA Change
Number of Clusters (i.e. Middle Schools)	480	480	476	274
Number of Cohort-Subject Observations	82,079	82,079	77,516	47,843
Number of Unique Students	584,449	584,449	580,905	399,640
Number of Student-Subject Test Scores	1,133,325	1,133,325	1,125,288	762,387

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

(b) Do Not Omit Years When Estimating VA

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.481	0.447	0.304	0.434
	(0.142)	(0.141)	(0.118)	(0.145)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Value Added Measure	Include All Years	Include All Years	Include All Years	Include All Years
Number of Clusters (i.e. Middle Schools)	480	480	476	291
Number of Cohort-Subject Observations	83,323	83,323	78,652	50,124
Number of Unique Students	587,918	587,918	584,300	404,041
Number of Student-Subject Test Scores	1,140,650	1,140,650	1,132,427	779,762

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

(c) Exclude Early Years

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.530	0.493	0.394	0.545
	(0.151)	(0.149)	(0.127)	(0.153)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Years	1998 - 2010	1998 - 2010	1998 - 2010	1998 - 2010
Number of Clusters (i.e. Middle Schools)	452	452	450	271
Number of Cohort-Subject Observations	68,816	68,816	65,268	45,468
Number of Unique Students	470,035	470,035	467,312	375,123
Number of Student-Subject Test Scores	910,372	910,372	904,174	715,169

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and

Table H2—Shrunken Leave-Out Mean Vs. Conventional Leave-Out Mean

(a) Elementary-to-Middle School Transition

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.512	0.505	0.530	0.414
	(0.146)	(0.146)	(0.151)	(0.119)
Type of Leave-Out Mean	Shrunken	Conventional	Shrunken	Conventional
Regression Type	2SLS	2SLS	OLS	OLS
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	479	479	480	480
Number of Cohort-Subject Observations	82,060	82,060	82,079	82,075
Number of Unique Students	584,280	584,280	584,449	584,427
Number of Student-Subject Test Scores	1,133,003	1,133,003	1,133,325	1,133,281

Each column reports coefficients from either a 2SLS or OLS regression. The columns differ in whether they use a shrunken leave-out mean or an conventional leave-out mean, as described in Appendix B. All the regressions are run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and all variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the school-year level.

(b) All Students

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.366	0.429	0.384	0.094
	(0.111)	(0.181)	(0.119)	(0.0465)
Type of Leave-Out Mean	Shrunken	Conventional	Shrunken	Conventional
Regression Type	2SLS	2SLS	OLS	OLS
Sample	All Students	All Students	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,266	1,266	1,278	1,277
Number of Cohort-Subject Observations	204,097	204,097	216,409	210,540
Number of Unique Students	1,236,014	1,236,014	1,256,986	1,250,657
Number of Student-Subject Test Scores	6,609,312	6,609,312	6,936,006	6,810,895

Each column reports coefficients from either a 2SLS or OLS regression. The columns differ in whether they use a shrunken leave-out mean or an conventional leave-out mean, as described in Appendix B. All regressions are run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and all variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the school-year level.

Table H3—Different Value-Added Measures

(a) Chetty, Friedman, Rockoff 2014 Controls

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.466	0.419	0.291	0.512
	(0.185)	(0.184)	(0.159)	(0.191)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Value Added Control Vector	Chetty, Friedman, and Rockoff 2014a	Chetty, Friedman, and Rockoff 2014a	Chetty, Friedman, and Rockoff 2014a	Chetty, Friedman, and Rockoff 2014a
Number of Clusters (i.e. Middle Schools)	480	480	476	274
Number of Cohort-Subject Observations	81,767	81,767	77,250	47,624
Number of Unique Students	583,767	583,767	580,247	399,259
Number of Student-Subject Test Scores	1,129,665	1,129,665	1,121,760	759,723

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

(b) No Within-Teacher VA Variation

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.470	0.448	0.339	0.476
	(0.175)	(0.174)	(0.148)	(0.178)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Value Added Measure	No Drift	No Drift	No Drift	No Drift
Number of Clusters (i.e. Middle Schools)	480	480	476	274
Number of Cohort-Subject Observations	81,961	81,961	77,435	47,818
Number of Unique Students	584,043	584,043	580,546	399,503
Number of Student-Subject Test Scores	1,132,569	1,132,569	1,124,613	762,127

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H4—IV Regressions

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.512	0.476	0.380	0.533
	(0.146)	(0.145)	(0.123)	(0.150)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	479	479	476	274
Number of Cohort-Subject Observations	82,060	82,060	77,500	47,843
Number of Unique Students	584,280	584,280	580,738	399,640
Number of Student-Subject Test Scores	1,133,003	1,133,003	1,124,969	762,387

Each column reports coefficients from a 2SLS regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. I use the measure described in Section III as an instrument for the overall average of the peers' previous teacher VA. The instrument is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. See Appendix C.2 for more information about the specification. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H5—Pseudo-Zoned Specification

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.850 (0.180)	0.782 (0.178)	0.578 (0.154)	0.775 (0.182)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	485	480	476	274
Number of Cohort-Subject Observations	101,872	82,079	77,516	47,843
Number of Unique Students	721,843	584,449	580,905	399,640
Number of Student-Subject Test Scores	1,403,812	1,133,325	1,125,288	762,387

Each column reports coefficients from a 2SLS regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. I use the measure described in Appendix C.3 as an instrument for the average of the peers' previous teacher VA. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H6—What is the Relevant Peer Group?

(a) Race

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Peers' Lagged Teacher VA - Same Peer Group	0.404	0.267	0.300	0.255
	(0.166)	(0.166)	(0.131)	(0.147)
Peers' Lagged Teacher VA - Other Peer Group	0.172	0.310	0.132	0.0974
	(0.175)	(0.166)	(0.130)	(0.140)
Probability Coefficients Are Equal	0.444	0.884	0.454	0.532
Own Previous Teacher VA		X		X
Current Teacher VA		X		X
Own Baseline Test Score		X		X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Peer Group Definition	Race	Race	Race	Race
Number of Clusters (i.e. Middle Schools)	465	269	1,247	1,034
Number of Cohort-Subject Observations	95,988	59,195	302,477	189,588
Number of Unique Students	514,814	355,264	1,236,681	1,099,411
Number of Student-Subject Test Scores	994,411	675,123	6,635,794	5,237,747

Each column reports coefficients from a regression run at the cohort-demographic group-subject level and weighted by the number of students in the cohort-demographic group combo who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and demographic groups are defined by an individual's race. The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. For each demographic group, I construct two measures which calculate the previous teacher quality of the teachers who taught both individuals who are in the same demographic group and those in the different demographic group. Their construction is discussed in detail in Section II, Section V, and Appendix C.4. All variables are constructed as the year-to-year change between two cohort-demographic groups who attended the same elementary and middle school, were in the same demographic group, and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

(b) Gender Specific VA Estimates

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Peers' Lagged Teacher VA - Same Peer Group	0.293	0.221	0.209	0.170
	(0.155)	(0.151)	(0.129)	(0.136)
Peers' Lagged Teacher VA - Other Peer Group	0.248	0.338	0.207	0.193
	(0.155)	(0.161)	(0.123)	(0.127)
Probability Coefficients Are Equal	0.858	0.651	0.994	0.914
Own Previous Teacher VA		X		X
Current Teacher VA		X		X
Own Baseline Test Score		X		X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Peer Group Definition	Gender	Gender	Gender	Gender
Number of Clusters (i.e. Middle Schools)	473	273	1,262	1,037
Number of Cohort-Subject Observations	105,727	64,099	244,938	167,844
Number of Unique Students	558,478	384,583	1,195,512	1,101,672
Number of Student-Subject Test Scores	1,081,244	732,073	6,204,589	5,324,073

Each column reports coefficients from a regression run at the cohort-demographic group-subject level and weighted by the number of students in the cohort-demographic group combo who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and demographic groups are defined by an individual's gender. The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. For each demographic group, I construct two measures which calculate the previous teacher quality of the teachers who taught both individuals who are in the same demographic group and those in the different demographic group. Their construction is discussed in detail in Section II, Section V, and Appendix C.4. All variables are constructed as the year-to-year change between two cohort-demographic groups who attended the same elementary and middle school, were in the same demographic group, and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

(c) English Language Learner Specific VA Estimates

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Peers' Lagged Teacher VA - Same Peer Group	0.550	0.647	0.401	0.463
	(0.202)	(0.192)	(0.184)	(0.179)
Peers' Lagged Teacher VA - Other Peer Group	0.435	0.195	0.527	0.383
	(0.377)	(0.391)	(0.281)	(0.285)
Probability Coefficients Are Equal	0.792	0.312	0.717	0.821
Own Previous Teacher VA		X		X
Current Teacher VA		X		X
Own Baseline Test Score		X		X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Peer Group Definition	English Language Learner	English Language Learner	English Language Learner	English Language Learner
Number of Clusters (i.e. Middle Schools)	384	237	1,080	920
Number of Cohort-Subject Observations	22,474	14,392	72,745	52,722
Number of Unique Students	245,512	183,517	896,993	849,764
Number of Student-Subject Test Scores	462,850	339,635	3,548,035	3,237,749

Each column reports coefficients from a regression run at the cohort-demographic group-subject level and weighted by the number of students in the cohort-demographic group combo who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and demographic groups are defined by hether an individual is classified as an English Language Learner (ELL). The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. For each demographic group, I construct two measures which calculate the previous teacher quality of the teachers who taught both individuals who are in the same demographic group and those in the different demographic group. Their construction is discussed in detail in Section II, Section V, and Appendix C.4. All variables are constructed as the year-to-year change between two cohort-demographic groups who attended the same elementary and middle school, were in the same demographic group, and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H7—Dynamic Spillovers

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(2) Ave. Test Score	(2) Ave. Test Score
Average of Peers' Previous Teacher VA	0.530	0.555	0.524	0.515
	(0.151)	(0.149)	(0.148)	(0.148)
Average of Peers' Twice Previous Teacher VA		0.350		0.290
		(0.176)		(0.173)
Own Previous Teacher VA			X	X
Own Twice Previous Teacher VA				X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition	Elem to Middle School Transition
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	480	478	478	478
Number of Cohort-Subject Observations	82,079	80,253	80,253	80,253
Number of Unique Students	584,449	568,440	568,440	568,440
Number of Student-Subject Test Scores	1,133,325	1,101,840	1,101,840	1,101,840

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The covariate "Ave. of Peers' Previous Teacher VA" is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. The covariate "Ave. of Peers' Twice Previous Teacher VA is constructed the same way, but uses the peers' twice previous teachers instead of their previous teachers. See Appendix C.5 for more details. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H8—Indirect Effect Estimates - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.384 (0.119)	0.343 (0.118)	0.289 (0.115)	0.358 (0.133)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,278	1,278	1,263	1,042
Number of Cohort-Subject Observations	216,409	216,409	171,193	124,373
Number of Unique Students	1,256,986	1,256,986	1,208,691	1,136,103
Number of Student-Subject Test Scores	6,936,006	6,936,006	6,345,198	5,712,727

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H9—Indirect Effect Estimates By Subject - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.373	0.350	0.400	0.371
	(0.131)	(0.147)	(0.170)	(0.195)
Own Average Previous Teacher VA		X		X
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math	Math	English	English
Number of Clusters (i.e. Schools)	1,276	1,038	1,271	1,027
Number of Cohort-Subject Observations	110,818	64,166	105,591	60,207
Number of Unique Students	1,246,428	1,124,625	1,215,290	1,088,460
Number of Student-Subject Test Scores	3,561,120	2,952,012	3,374,886	2,760,715

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H10—Falsification Test - Full Sample

	(1) Ave. of Peers' Previous Teacher VA	(2) Ave. of Peers' Previous Teacher VA	(3) Ave. of Peers' Previous Teacher VA	(4) Ave. of Peers' Previous Teacher VA	(5) Ave. of Peers' Previous Teacher VA	(6) Ave. of Peers' Previous Teacher VA	(7) Ave. of Peers' Previous Teacher VA
Baseline Test Scores	0.0000942						0.0000631
	(0.0000590)						(0.0000654)
Percent English Language Learner		0.000323					0.000174
		(0.000356)					(0.000498)
Percent on Free or Reduced Lunch			0.00032				0.000295
			(0.000142)				(0.000153)
Percent Black				-0.000288			-0.000852
				(0.000251)			(0.000561)
Percent Hispanic					-0.000112		-0.000667
					(0.000221)		(0.000499)
Percent White						0.000275	-0.0000959
						(0.000348)	(0.000607)
Sample	All Schools	All Schools	All Schools	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,263	1,278	1,238	1,277	1,277	1,277	1,222
Number of Cohort-Subject Observations	171,193	216,409	166,579	216,378	216,378	216,378	134,048
Number of Unique Students	1,208,691	1,256,986	1,012,819	1,256,975	1,256,975	1,256,975	975,646
Number of Student-Subject Test Scores	6,345,198	6,936,006	5,438,683	6,935,975	6,935,975	6,935,975	5,061,813

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned form the same elementary school to the same middle school in the same year. The outcome measure, "Ave. of Peers' Previous Teacher VA," is designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H11—Effect of Past and Future Teacher Transitions - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.521	0.465	0.395	0.479
	(0.157)	(0.157)	(0.151)	(0.175)
Twice Lag Ave. of Peers' Previous Teacher VA	-0.115	-0.142	-0.155	-0.126
	(0.151)	(0.153)	(0.144)	(0.158)
Lag Ave. of Peers' Previous Teacher VA	-0.127	-0.193	-0.110	-0.210
	(0.157)	(0.158)	(0.143)	(0.159)
Lead Ave. of Peers' Previous Teacher VA	0.225	0.168	0.136	0.215
	(0.144)	(0.144)	(0.136)	(0.162)
Twice Lead Ave. of Peers' Previous Teacher VA	-0.0133	-0.0300	-0.0238	-0.0357
	(0.139)	(0.138)	(0.127)	(0.142)
Probability All Five Above Coefficients Are Equal	0.0210	0.0300	0.0810	0.0300
Probability All Four Lag and Lead Coefficients Are Zero	0.361	0.346	0.506	0.314
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,056	1,056	1,052	928
Number of Cohort-Subject Observations	78,917	78,917	69,497	56,511
Number of Unique Students	916,887	916,887	888,107	847,799
Number of Student-Subject Test Scores	4,403,134	4,403,134	4,083,105	3,808,054

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariate, "Ave. of Peers' Previous Teacher VA," as well as all of its leads and lags are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Its construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H12—Effect of Future Peers - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. Of Future Peers' Previous Teacher VA	0.187	-0.0377		
	(0.105)	(0.0829)		
Ave. Of Future Peers' Current Teacher VA			0.134	0.0493
			(0.116)	(0.0912)
Own Average Previous Teacher VA		X		X
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,235	1,027	1,034	1,026
Number of Cohort-Subject Observations	266,102	214,201	237,340	198,914
Number of Unique Students	1,019,514	951,488	1,022,350	912,946
Number of Student-Subject Test Scores	4,747,533	4,228,376	4,576,329	3,864,293

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. Both of the covariates are designed to capture the teacher quality at the schools that feed the students' future school, but which he or she did not attend and described in detail in Section III.B.2 and Appendix E. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H13—Do Spillovers Occur Within Subjects? - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. of Peers' Previous Teacher VA - Same Subject	0.341	0.297	0.257	0.300
	(0.100)	(0.1000)	(0.102)	(0.114)
Ave. of Peers' Previous Teacher VA - Other Subject	0.0674	0.0710	0.0491	0.0993
	(0.105)	(0.105)	(0.103)	(0.128)
Probability Coefficients Are Equal	0.0520	0.106	0.156	0.234
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,277	1,277	1,262	1,041
Number of Cohort-Subject Observations	213,310	213,310	169,222	122,737
Number of Unique Students	1,255,023	1,255,023	1,206,827	1,134,012
Number of Student-Subject Test Scores	6,903,365	6,903,365	6,313,985	5,688,893

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Their construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H14—Do Spillovers Occur Within Subjects? - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. of Peers' Previous Teacher VA - Same Subject	0.343	0.299	0.336	0.304
	(0.151)	(0.159)	(0.195)	(0.221)
Ave. of Peers' Previous Teacher VA - Other Subject	0.0444	0.0991	0.0843	0.0973
	(0.180)	(0.200)	(0.162)	(0.193)
Probability Coefficients Are Equal	0.295	0.507	0.416	0.563
Own Average Previous Teacher VA		X		X
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math	Math	English	English
Number of Clusters (i.e. Schools)	1,275	1,037	1,271	1,027
Number of Cohort-Subject Observations	109,161	63,314	104,149	59,423
Number of Unique Students	1,245,176	1,123,478	1,214,289	1,087,281
Number of Student-Subject Test Scores	3,543,706	2,939,417	3,359,659	2,749,476

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. Their construction is discussed in detail in Section II. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H15—What is the Relevant Peer Group? - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. of Peers' Lagged Teacher VA - Same Race and Gender	0.305	0.287	0.159	0.228
	(0.107)	(0.107)	(0.0959)	(0.114)
Ave. of Peers' Lagged Teacher VA - Same Race and Different Gender	0.144	0.132	0.185	0.224
	(0.105)	(0.105)	(0.0911)	(0.106)
Ave. of Peers' Lagged Teacher VA - Different Race and Same Gender	0.0444	0.0251	0.0285	0.0171
	(0.0909)	(0.0900)	(0.0801)	(0.0903)
Ave. of Peers' Lagged Teacher VA - Different Race and Gender	-0.0230	-0.0347	-0.0719	-0.0552
	(0.0878)	(0.0870)	(0.0847)	(0.0952)
Probability Same Race/Gender Coefficient Is Equal to Different Race/Gender Coefficient	0.0350	0.0370	0.115	0.0970
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,215	1,215	1,212	1,025
Number of Cohort-Demographic Group-Subject Observations	383,604	383,604	305,244	253,216
Number of Unique Students	1,182,759	1,182,759	1,101,988	1,037,904
Number of Student-Subject Test Scores	6,039,750	6,039,750	5,099,435	4,639,254

Each column reports coefficients from a regression run at the cohort-demographic group-subject level and weighted by the number of students in the cohort-demographic group combo who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and demographic groups are defined as those by an individual's gender and race. The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. For each demographic group, I construct four measures which calculate the previous teacher quality of the teachers who taught four groups of students: individuals who are the same race and gender as the demographic group, individuals who are the same race but a different gender, individuals who are a different gender but the same race, and individuals who are both a different race and gender. Their construction is discussed in detail in Section II and Section V. All variables are constructed as the year-to-year change between two cohort-demographic groups who attended the same two schools in a row, were in the same demographic group, and who made the transition between schools in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H16—What is the Relevant Peer Group? - Full Sample

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. of Peers' Lagged Teacher VA - Same Peer Group	0.367	0.383	0.308**	0.201
	(0.108)	(0.124)	(0.140)	(0.157)
Ave. of Peers' Lagged Teacher VA - Other Peer Group	0.0475	-0.104	0.123	0.201
	(0.108)	(0.120)	(0.143)	(0.163)
Probability Coefficients Are Equal	0.0650	0.0190	0.457	1
Own Average Previous Teacher VA		X		X
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	All Schools	All Schools	All Schools	All Schools
Subjects	Math and English	Math and English	Math and English	Math and English
Peer Group Definition	Race	Race	Gender	Gender
Number of Clusters (i.e. Schools)	1,244	1,032	1,242	1,035
Number of Cohort-Demographic Group-Subject Observations	289,344	183,176	244,253	157,315
Number of Unique Students	1,225,977	1,090,966	1,244,399	1,107,314
Number of Student-Subject Test Scores	6,520,749	5,164,108	6,735,433	5,352,737

Each column reports coefficients from a regression run at the cohort-demographic group-subject level and weighted by the number of students in the cohort-demographic group combo who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year and demographic groups are defined either by an individual's race (in columns (1) and (2)) or gender (in columns (3) and (4)). The main covariates are designed to capture the teacher quality at the elementary schools that feed the students' middle school, but which they did not attend. For each demographic group, I construct two measures which calculate the previous teacher quality of the teachers who taught both individuals who are in the same demographic group and those in the different demographic group. Their construction is discussed in detail in Section II and Section V. All variables are constructed as the year-to-year change between two cohort-demographic groups who attended the same two schools in a row, were in the same demographic group, and who made the transition between schools in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H17—First Specification Test

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Leave-Out Average Peers' Previous Teacher VA	-0.0847	-0.204	-0.0229	-0.0314
	(0.139)	(0.139)	(0.0548)	(0.0567)
Leave-Out Average Peers' Previous Teacher VA x Fraction of Peers Affected	0.567	0.777	0.399	0.433
	(0.239)	(0.242)	(0.142)	(0.158)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Students	All Students	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	853	633	1,277	1,041
Number of Cohort-Subject Observations	86,866	52,530	210,540	122,909
Number of Unique Students	745,920	553,514	1,250,657	1,127,059
Number of Student-Subject Test Scores	1,465,043	1,071,233	6,810,895	5,623,093

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. "Leave-Out Average Peers' Previous Teacher VA" is the leave-out average Teacher VA at the schools that feed a student's current school, but which he or she did not attend. "Fraction of Peers" corresponds to the fraction of students at the individual's current school, who previously attended a different school. For more details, see Appendix D.2. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H18—Second Specification Test

VARIABLES	(1) Test Score	(2) Test Score	(3) Test Score	(4) Test Score
Average Own Previous Teacher VA	0.205	0.136	0.305	0.267
	(0.0591)	(0.0586)	(0.0423)	(0.0437)
Average Own Previous Teacher VA x Fraction of Peers Affected	0.667	0.786	0.146	0.152
	(0.329)	(0.322)	(0.0574)	(0.0581)
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	Elem to Middle School Transition	Elem to Middle School Transition	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	480	274	1,278	1,042
Number of Cohort-Subject Observations	82,079	47,843	216,409	124,373
Number of Unique Students	584,449	399,640	1,256,986	1,136,103
Number of Student-Subject Test Scores	1,133,325	762,387	6,936,006	5,712,727

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. "Fraction of Peers Affected" corresponds to the fraction of students at the individual's current school, who previously attended the same school. For more details, see Appendix D.2. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H19—Combining the Falsification and Specification Tests

	(1) Baseline Test Scores	(2) Percent English Language Learner	(3) Percent on Free or Reduced Lunch	(4) Percent Black	(5) Percent Hispanic	(6) Percent White
Average of Peers' Previous Teacher VA	-0.0471	0.0291	0.0392	-0.0230	0.0142	-0.00373
	(0.179)	(0.0547)	(0.141)	(0.0697)	(0.0716)	(0.0438)
Leave-Out Average of Peers' Previous Teacher VA	0.118	-0.0183	0.184	0.00821	-0.0199	-0.0210
	(0.0954)	(0.0326)	(0.0732)	(0.0326)	(0.0337)	(0.0189)
Fraction of Peers Affected	-0.00455	-0.000352	-0.000522	-0.000930	0.000202	0.00120
	(0.00322)	(0.000890)	(0.00316)	(0.00145)	(0.00150)	(0.00108)
Sample	All Students	All Students	All Students	All Students	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,263	1,277	1,236	1,276	1,276	1,276
Number of Cohort-Subject Observations	169,522	210,540	161,855	210,509	210,509	210,509
Number of Unique Students	1,202,056	1,250,657	1,006,932	1,250,646	1,250,646	1,250,646
Number of Student-Subject Test Scores	6,246,459	6,810,895	5,332,306	6,810,864	6,810,864	6,810,864

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned form the same elementary school to the same middle school in the same year. "Average of Peers' Previous Teacher VA" is the average Teacher VA at the schools that feed a student's current school, but which he or she did not. "Leave-Out Ave. of Peers' Previous Teacher VA" is the same variable, but not weighted by the fraction of peers who attended other elementary schools. See Appendix D.2 for more discussion about the differences between the two variables. "Fraction of Peers Affected" corresponds to the fraction of students at the individual's current school, who previously attended a different school. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H20—Combining the First Placebo and Specification Tests

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Average of Peers' Previous Teacher VA	0.523	0.496	0.504	0.571
	(0.194)	(0.194)	(0.188)	(0.212)
Twice Lag Ave. of Peers' Previous Teacher VA	-0.241	-0.250	-0.362	-0.256
	(0.190)	(0.191)	(0.181)	(0.197)
Lag Peers' Ave. of Previous Teacher VA	-0.228	-0.265	-0.229	-0.308
	(0.201)	(0.203)	(0.188)	(0.206)
Lead Peers' Ave. of Previous Teacher VA	0.126	0.104	0.0351	0.175
	(0.181)	(0.182)	(0.174)	(0.200)
Twice Lead Ave. of Peers' Previous Teacher VA	0.113	0.116	0.150	0.143
	(0.188)	(0.187)	(0.173)	(0.189)
Twice Lag Leave-Out Ave. of Peers' Previous Teacher VA	0.115	0.103	0.168	0.117
	(0.0768)	(0.0764)	(0.0768)	(0.0788)
Lag Leave-Out Ave. of Peers' Previous Teacher VA	0.0887	0.0703	0.0974	0.0847
	(0.0773)	(0.0768)	(0.0783)	(0.0781)
Leave-Out Ave. of Peers' Previous Teacher VA	0.0341	0.0156	-0.0437	-0.0263
	(0.0739)	(0.0738)	(0.0760)	(0.0753)
Lead Leave-Out Ave. of Peers' Previous Teacher VA	0.0792	0.0574	0.0877	0.0496
	(0.0699)	(0.0696)	(0.0710)	(0.0724)
Twice Lead Leave-Out Ave. of Peers' Previous Teacher VA	-0.0479	-0.0594	-0.0877	-0.0873
	(0.0737)	(0.0736)	(0.0729)	(0.0736)
Twice Lag Fraction of Peers Affected	0.059	0.0556	0.0412	0.0395
	(0.0271)	(0.0268)	(0.0254)	(0.0267)
Lag Fraction of Peers Affected	-0.0293	-0.0240	-0.0319	-0.00930
	(0.0432)	(0.0431)	(0.0406)	(0.0431)
Fraction of Peers Affected	-0.00832	-0.0101	0.0102	-0.0122
	(0.0415)	(0.0412)	(0.0395)	(0.0413)
Lead Fraction of Peers Affected	0.0300	0.0295	0.0152	0.00919
	(0.0335)	(0.0335)	(0.0318)	(0.0342)
Twice Lead Fraction of Peers Affected	-0.0585	-0.0585	-0.0418	-0.0338
	(0.0250)	(0.0247)	(0.0242)	(0.0250)
Own Average Previous Teacher VA		X	X	X
Own Average Baseline Test Score			X	X
Current Average Teacher VA				X
Sample	All Students	All Students	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	1,042	1,042	1,040	913
Number of Cohort-Subject Observations	72,728	72,728	66,838	54,044
Number of Unique Students	898,560	898,560	869,974	827,071
Number of Student-Subject Test Scores	4,199,169	4,199,169	3,919,627	3,654,144

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. "Average of Peers' Previous Teacher VA" is the average Teacher VA at the schools that feed a student's current school, but which he or she did not. "Leave-Out Ave. of Peers' Previous Teacher VA" is the same variable, but not weighted by the fraction of peers who attended other elementary schools. See Appendix D.2 for more discussion about the differences between the two variables. "Fraction of Peers Affected" corresponds to the fraction of students at the individual's current school, who previously attended a different school. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

Table H21—Combining the Second Placebo and Specification Tests

VARIABLES	(1) Ave. Test Score	(2) Ave. Test Score	(3) Ave. Test Score	(4) Ave. Test Score
Ave. Of Future Peers' Previous Teacher VA	0.102	-0.114		
	(0.174)	(0.142)		
Leave-Out Ave. Of Future Peers' Previous Teacher VA	-0.00173	0.00283		
	(0.00316)	(0.00199)		
Ave. Of Future Peers' Current Teacher VA			0.133	0.0593
			(0.146)	(0.126)
Leave-Out Ave. Of Future Peers' Current Teacher VA			-0.000587	0.00311
			(0.00267)	(0.00174)
Fraction of Peers from Other Elementary	0.0641	0.0546	0.00908	-0.00852
	(0.113)	(0.0921)	(0.0637)	(0.0646)
Own Average Previous Teacher VA		X		X
Own Average Baseline Test Score		X		X
Current Average Teacher VA		X		X
Sample	All Students	All Students	All Students	All Students
Subjects	Math and English	Math and English	Math and English	Math and English
Number of Clusters (i.e. Middle Schools)	1,234	1,027	1,031	1,024
Number of Cohort-Subject Observations	261,238	209,972	226,584	194,608
Number of Unique Students	1,016,920	949,100	1,010,752	906,509
Number of Student-Subject Test Scores	4,715,611	4,199,842	4,455,882	3,793,516

Each column reports coefficients from a regression run at the cohort-subject level and weighted by the number of students in the cohort who took the test in the relevant subject. A cohort is defined as a group of students who transitioned from the same elementary school to the same middle school in the same year. "Average of Future Peers' Previous Teacher VA" is the average previous teacher VA at the schools that feed a student's future middle school school, but which he or she did not. "Leave-Out Ave. of Future Peers' Previous Teacher VA" is the same variable, but not weighted by the fraction of peers who attended other elementary schools. See Appendix D.2 for more discussion about the differences between the two variables. "Ave. of Future Peers' Current Teacher VA" and "Leave-Out Ave. Of Future Peers' Current Teacher VA" are defined similarly. "Fraction of Peers Affected" corresponds to the fraction of students at the individual's future middle school, who currently attend a different elementary school. All variables are constructed as the year-to-year change between two cohorts who attended the same elementary and middle school and who made the transition in subsequent years. Standard errors, in parenthesis, are clustered at the middle school level. The baseline test scores correspond to the test scores two-years prior to the current test score.

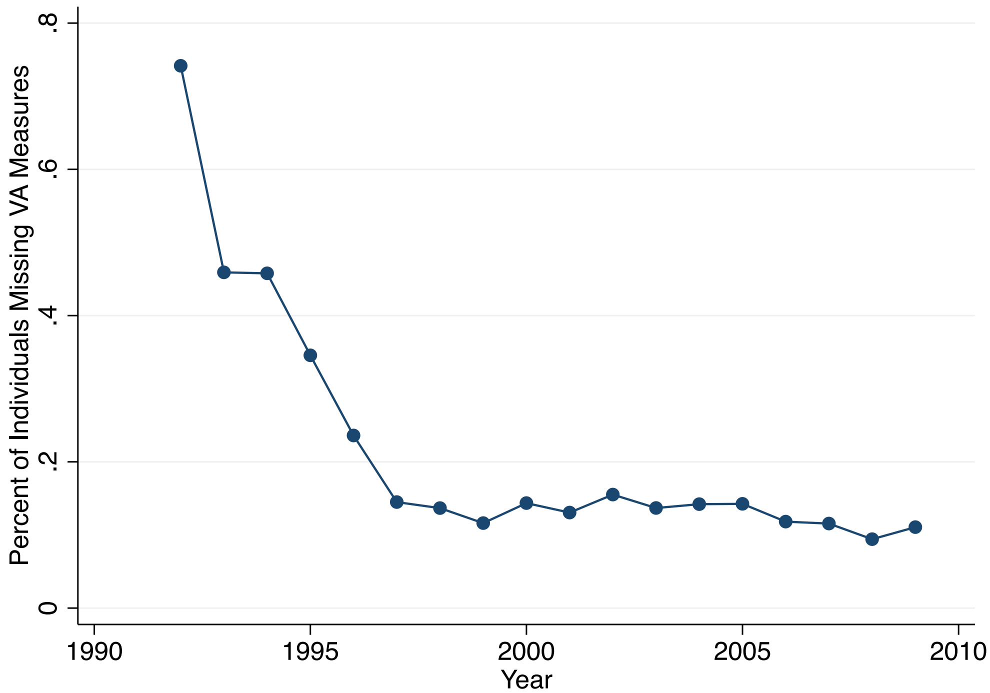
Table H22—Fade-Out of Direct Teacher Effect

VARIABLES	(1) Test Score	(2) Next Year's Test Score
Teacher VA	0.977	0.559
	(0.0369)	(0.0475)
Sample	All Students	All Students
Subjects	Math and English	Math and English
Number of Clusters (i.e. Schools)	1,067	1,058
Number of Cohort-Subject Observations	59,770	53,274
Number of Unique Students	1,384,145	1,384,145
Number of Student-Subject Test Scores	8,060,148	8,060,148

Each column reports coefficients from an OLS regression that regresses changes in cohort test scores on changes in cohort teacher value added. A cohort here is defined as a school-grade-subject-year cell. Standard errors, in parenthesis, are clustered at the school level, and all regressions are weighted by the number of students in the cohort.

H2. Appendix Figures

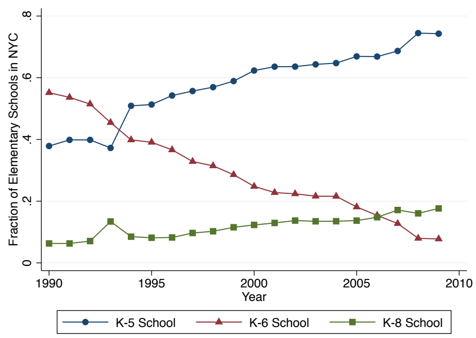
Figure H1. Fraction of 6th Grade Students Missing 5th Grade Teacher Value-Added Estimates



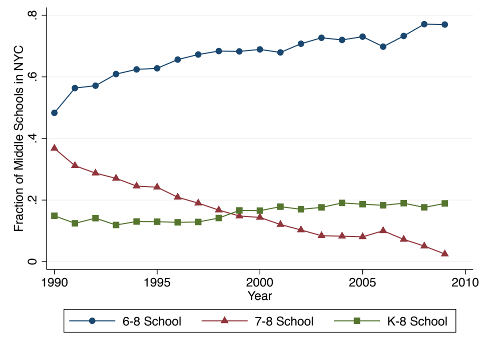
Note: This figure shows the fraction of first year middle school students who are missing a teacher VA estimate for their previous math and English teacher. As is clear, few individuals are matched to a teacher VA in the early years of the data, but the match quickly rate increases and stabilizes at a little over 90%. It never reaches 100% matches for two reasons. First, any student who is new to the New York City public school system will not be matched to a previous teacher. Second, I cannot estimate VA for teachers who teach for fewer than three years in New York City, because I exclude two years of data from the estimation.

Figure H2. Elementary-to-Middle School Transitions in New York City

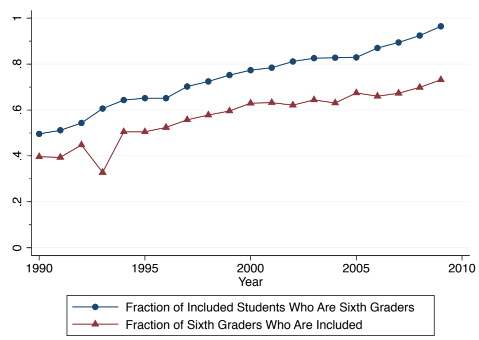
(a) Elementary Schools in New York City



(b) Middle Schools in New York City



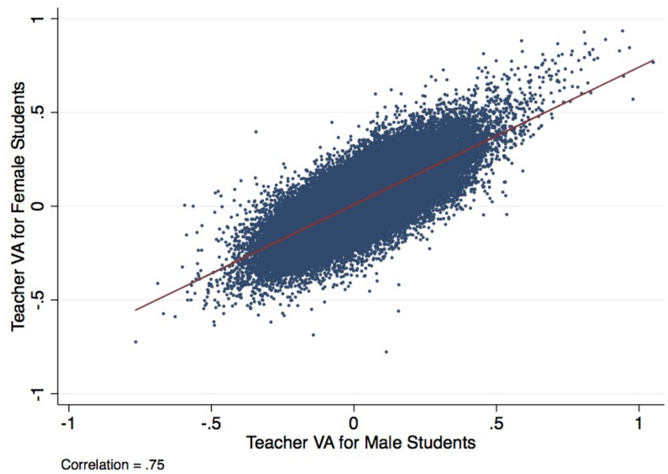
(c) Included Students in Sample



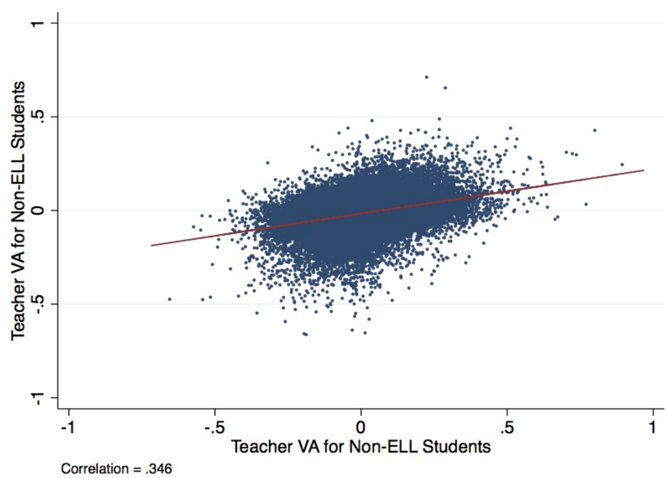
Note: These figures illustrate how the structure of elementary and middle schools changed over the time period. Figure H2a shows that the fraction of 5^{th} grade students in New York City who attended a K-6 elementary school decreases dramatically over the time period, as more students began to attend K-5 and K-8 elementary schools. Figure H2b illustrates the reverse trend for middle schools; the fraction of 8^{th} grade students who attended a middle school serving only grades 7 and 8 decreased, as more schools began serving grades 6 through 8 or K through 8. Together these trends imply both that the fraction of 6^{th} graders who are included in the sample increases over the time period, as does the fraction of students in the sample who are 6^{th} graders instead of 7^{th} graders. This is shown in Figure H2c.

Figure H3. Correlations Between Subgroup Specific Value-Added

(a) Gender Specific Value-Added Estimates

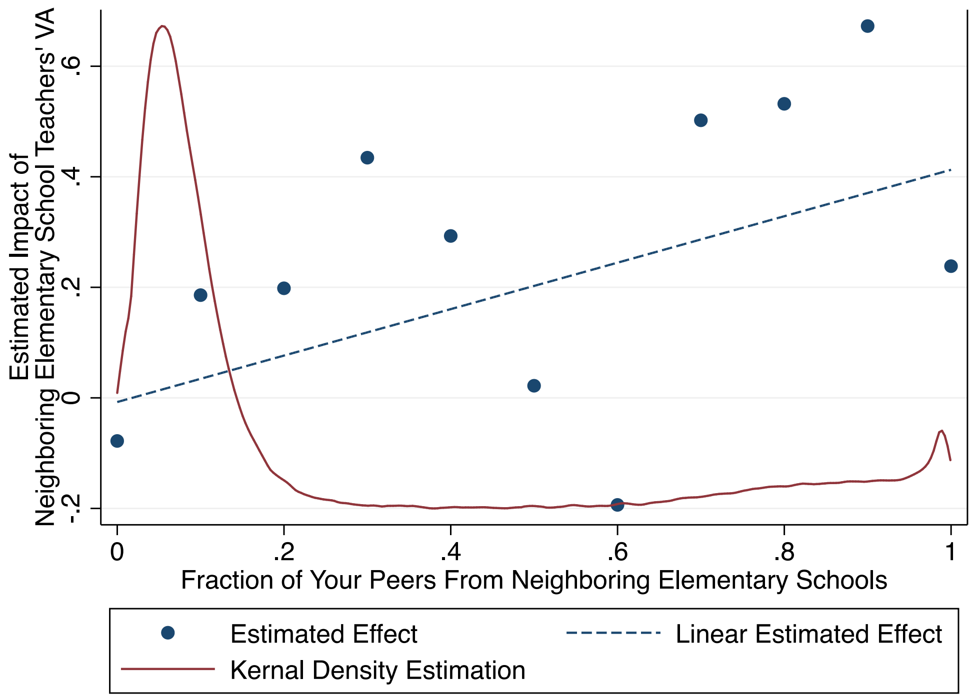


(b) ELL Specific Value-Added Estimates



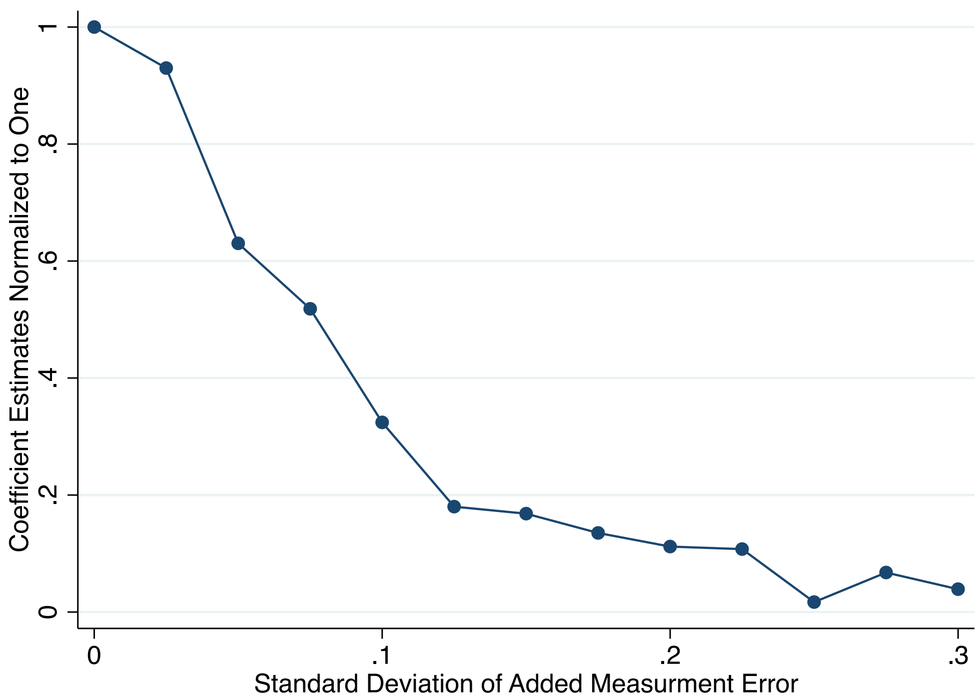
Note: These figures show the within-teacher-year correlation between different teacher VA measures. The top figure shows how a teacher's VA for students who are males is correlated with the same teacher's VA for students who are female; the bottom figure shows how a teacher's VA for students who are classified as an English Language Leaner (ELL) is correlated with the same teacher's VA measures of those who are not. The red line shows the estimated linear relationship between the two measures. For both the regression that gave rise to the red line and the estimated correlation, I weight each teacher using the number of students who each teacher taught.

Figure H4. Specification Test



Note: The above figure shows the estimated coefficients from Equation (D3), as well as the line implied from the linear specification. In addition, the solid red line shows the distribution of the fraction of a student's current peers who previously attended a different school than the student. The bimodality of this distribution implies that the coefficient estimates on the extremes of the x-axis are estimated more precisely than those in the middle.

Figure H5. Measurement Error Simulation



Note: The above figure shows the results of a simulation, where additional measurement error was first added to the teacher VA estimates and then the aggregation and estimation was conducted in the same manner as is done in the main analysis. The x-axis shows the standard deviation of the added measurement error, while the y-axis shows the estimated coefficient relative to the estimated coefficient when there is no added measurement error. The graph shows that despite the fact that measurement error can sometimes increase the point estimates of peer effects as discussed in Feld and Zölitz (2017), here additional measurement error in the teacher VA estimates causes the estimated peer effect to attenuate.