The Effects of Making Performance Information Public: Regression Discontinuity Evidence from Los Angeles Teachers*

Peter Bergman† and Matthew J. Hill‡

This paper uses school-district data and a regression discontinuity design to study the effects of making teachers' value-added ratings available to the public and searchable by name. We find that classroom compositions change as a result of this new information. In particular, high-scoring students sort into the classrooms of published, high-value added teachers. This sorting occurs when there is within school-grade variation in teachers' value added.

JEL Codes: I28, J08, J24.

*We thank Leah Platt Boustan, Richard Buddin, Dora Costa, Scott Imberman, Mike Lovenheim, Steve Rivkin, Jonah Rockoff, Eric Taylor, Gema Zamarro and several anonymous referees for their detailed feedback, as well as seminar participants at AEFP, CESifo, Toulouse and UCLA. All errors are our own.

†Teachers College, Columbia University; E-mail: *bergman@tc.columbia.edu*

‡Loyola Marymount University ; E-mail: *matthew.hill@lmu.edu*

1 Introduction

Performance information is publicly available in a variety of settings. Increasingly, and perhaps most controversially, performance information has been made public in the education sector. Many school districts, including the two largest districts in the country, New York City and Los Angeles, make school report cards available online, and the non-profit *GreatSchools* has formulated and published ratings online for more than 200,000 schools across the country.[1] In a number of places, including districts in Florida, Cleveland, New York and Los Angeles, individual teacher performance information has also been made public. The impacts of making this information public in the education context are of particular importance because researchers have quantified the significant social and economic value generated by high-quality teachers and schools (Rockoff 2004; Deming 2011; Hanushek 2011; Jackson 2013; Chetty et al. 2014).

This paper uses a discontinuity in the publication of teacher ratings in Los Angeles Unified School District (LAUSD) to study the effects of making performance information public on teacher retention, student sorting, and test scores. In August of 2010, the *Los Angeles Times* (*LA Times*) published teacher ratings for third through fifth grade teachers in Los Angeles Unified School District (LAUSD). These ratings were based on the newspaper's calculation of teachers' value added. An independent contractor for the *LA Times* computed these scores for English and math using student-level regressions.

That year, the *LA Times* heralded its ratings with a front-page article and provided free access to the value-added scores and ratings via an online, searchable database. The subsequent year, 2011, the *LA Times* published scores based on data through spring 2010 for almost all third-, fourth-, and fifth-grade teachers (almost) irrespective of the number of students the teacher had taught previously.[2] However, in the initial year, only teachers who had taught at least 60 students with test scores and lagged test scores between spring

[1] See www.greatschools.org.

[2] The calculations excluded teachers who had taught ten students or fewer.

2003 and spring 2009 were published in the *LA Times*. We use this threshold to identify the effects of ratings publication on classroom outcomes.

The release of this information could affect how students sort into teachers' classrooms. Both parents and teachers can influence this allocation, and previous research shows parents' demand for higher test scores (Black 1999; Figlio and Lucas 2004). Principals may adhere to parent requests for their children to be placed with preferred teachers (Hui 2003). Experienced teachers may also be able to successfully lobby the administration for good students or to have "problem" students placed elsewhere (Pop-Eleches and Urquiola 2013). Clotfelter et al. (2006) suggest that, in aggregate, more able students are more likely to be placed with high-quality teachers and Jacob and Lefgren (2007) found that even low-income parents petition schools for their children to have higher value-added teachers.

Publishing teacher ratings could also affect teacher retention within schools and the district. Bates (2016) showed how the release of value-added information could reduce informational asymmetries between principals and teachers, promoting within-district mobility for high value-added teachers and out-of-district mobility for low value-added teachers. Cullen et al. (2017) similarly found that the informing principals in Houston about teacher effectiveness increases the rate of exit from the district by lower-rated teachers. Rockoff et al. (2012) found that principals update their beliefs about teachers' performance when provided information on their value-added. This occurs despite principals' strong prior beliefs about teachers' value-added that correlate with estimated value-added. Rockoff et al. (2012) also found that subsequently low value-added teachers were more likely to exit the school district. Lastly, Adnot et al. (2017) showed that, while teacher turnover correlates negatively with student performance, a teacher performance-assessment and incentive program helped District of Columbia Public Schools identify and replace low-rated teachers, which led to increases in student performance.

We find that publication causes students with high test scores to sort into classrooms with high value-added teachers compared with similar teachers whose value-added score is *not* published. For instance, a one standard deviation increase in a teacher's value-added causes the math test scores of incoming students to be three tenths of a standard deviation higher on average than those of a similar teacher whose rating is not published. This finding is consistent with students and parents who value academic performance seeking out higher rated teachers, though schools could shift students across classrooms as well. We also find a similar observed effect when we use teacher quality ratings that are relative to teachers within the same grade level and school instead of overall teacher value-added. This suggests that sorting will exist as long there is variance among the set of possible teachers for students. Finally, we find evidence that low-rated teachers are more likely to exit the district as a result of publication compared with similar unpublished teachers.

In a paper related to ours, Imberman and Lovenheim (2016) examined whether housing markets responded to the publication of additional school performance information in Los Angeles by the *LA Times*. They found that this new information on schools' value-added is not capitalized into housing prices. However, parents may have responded to the publication of performance information in other ways. For instance, to the extent that parents can influence teacher assignment *within* a school academically oriented parents may push for their children to be assigned to high-rated teachers.

Lastly, Pope (2017) has concurrently written about the publication of teachers' value-added in Los Angeles, which was made available around the same time as our paper. Pope uses an event-study design to analyze the impacts of publication on students' test scores, while controlling for classroom composition (e.g. through students' prior test scores). He does not, however, find evidence of student sorting using this research design. Pope uses different variation than we do: he relies on across-time variation in classroom characteristics among the same group of teachers before and after publication. Our paper uses within-year variation in classroom composition. By comparing published to unpublished teachers within the same year, we estimate the net change between these teachers' classrooms. This implies our results reflect a different estimand than Pope's event-study analysis of published teachers. For instance, our findings potentially capture high-scoring students moving out of unpublished, highly-rated teachers' classrooms and into published, highly-rated teachers' classrooms. A related caveat to our analysis is that the sample of teachers for our regression-discontinuity estimates are slightly less experienced (14 years versus 12 years) than the overall sample, and less likely to be tenured (95% to 91%). This implies our estimates may not extrapolate to the set of all teachers, though our sample is nonetheless highly experienced and overwhelmingly tenured, like the overall sample.

The rest of this paper proceeds as follows. In Section 2 we provide background information on the release of teachers' value-added scores. In Section 3 we discuss the data and the empirical strategy. Sections 4 shows effects of publication, and Section 5 concludes.

2 Background

In August 2010, the *LA Times* published teacher value-added scores for third- through fifth-grade teachers in the Los Angeles Unified School District. The *LA Times* hired Dr. Richard Buddin, a senior economist at the RAND Corporation, to construct the scores. Details on Buddin's methodology can be found in his white paper on the subject (Buddin 2010). Buddin used methods commonly found in the literature: linear regression with teacher fixed effects controlling for student covariates (Jackson et al. 2014).[3] The value-added scores were based on student test score data from the 2002-2003 through 2008-2009 academic years obtained by the *LA times* via a Public Records Act request. A later release of the value-added scores was updated to include test score data from 2009-2010 as well.

The *LA Times* heralded its ratings with a front-page story and then provided the public with an online database of teachers and their corresponding value-added scores.[4] This database is searchable by teacher name and school. Access to the website is free, with no registration required. Figure 1 shows how the results are presented for a sample teacher.

[3]Buddin used similar value-added scores previously in a 2009 *Journal of Urban Economics* article coauthored with Gema Zamarro.

[4]http://projects.latimes.com/value-added/

The evaluation of a teacher consists of an overall score as well as a score for math and English. Scores are divided into five rating categories: "Least Effective," "Less Effective," "Average," "More Effective" and "Most Effective." These categories correspond to quintiles in the calculated value-added scores. The publication of the value-added scores was teachers' first exposure to numerical ratings as LAUSD had not previously computed scores of this type.

The publication of the value-added scores received widespread coverage, and there is substantial evidence that the public was aware of their release. The *LA Times* published 37 articles related to the value-added scores in the subsequent nine months following the initial release (Imberman and Lovenheim 2016). The scores were covered nationally by outlets such as the *New York Times*, *Washington Post*, *National Public Radio*, and *Fox News*. Locally, the scores received attention from both English- and Spanish-language news and radio stations, suggesting that knowledge of the scores extended across race and language barriers. The online database received over 230,000 page views on its first day (Song 2010). While there was widespread coverage of the scores, their publication upset many teachers (Lee 2011). Both the LAUSD teachers' union and the American Federation of Teachers criticized the *LA Times* for the release of the value-added scores. Teachers engaged in a series of protests against the *LA Times* culminating with a march on the *LA Times* building on September 14, 2010.

The initial release of the value-added scores was limited to teachers who had taught 60 or more tested students between the 2002-2003 and 2008-2009 academic years. Students needed to have at least one year with a test score and a lagged test score to be counted.[5] This 60-student cutoff provides a natural experiment. Teachers right below the cutoff should be similar to those just above the cutoff, which allows us to use a regression discontinuity design to estimate the impact of score publication on various outcomes.

[5]The 60-student number was chosen because of concerns from the *LA Times* that scores for those teachers with fewer than 60 students would be unreliable. (This is not standard practice.) The concerns about reliability proved to be unfounded and even if they were not, given our regression discontinuity design there is no reason to believe that expected precision or bias change discontinuously around the 60-student cutoff.

Figure 2 shows the timeline of the waves of release. Following the initial release in August 2010, in June 2011, the *LA Times* updated its value-added scores to include student test score data from the 2009-2010 school year (Buddin 2011). For the updated release, the *LA times* removed the 60-student cutoff rule and published value-added scores for all teachers who had at least 10 students fitting the criteria described above. In April 2011, LAUSD produced its own value-added measure, and the results were provided to teachers privately. These value-added scores were constructed by the Value-Added Research Center. The scores were calculated using a value-added methodology similar to the one used to generate published scores (Song and Felch 2011). [6] LAUSD denoted these scores as Academic Growth over Time (AGT) and provided them to teachers and principals in the spring of 2011. The release of these scores occurred after the time period used for the bulk of our empirical analysis, though this is important context for analyses of outcomes from the 2011-2012 school year.

Figure 2 also shows the relevant dates for the school year in the LAUSD. Most LAUSD schools are not open year-round; the academic year begins in September and ends in June. We find significant effects of value-added score publication on classroom composition; therefore, the timing of classroom assignment is relevant. Classroom assignment is an idiosyncratic process; assignments are made over the summer, but there is no standard time period within the summer during which schools assign students to teachers. Based on conversations with educators in LAUSD, we have marked the possible assignment period on the timeline.

3 Data and Empirical Strategy

The data used in the paper comes from three sources: LAUSD, the *LA Times* and the Common Core of Data. We filed a public records request with LAUSD to obtain similar data to those which the *LA Times* used to calculate the value-added scores: identifiable teachers linked to de-identified student test scores. We have these data for LAUSD students

[6] Imberman and Lovenheim (2016) found that the internal school-level scores were not highly correlated with the *LA Times* scores. The two scores use different time periods: the *LA Times* scores use data from the 2002-2003 to 2009-2010 academic years while the LAUSD internal scores used data only from the 2007-2008 to 2009-2010 academic years.

enrolled from the 2008-2009 school year through the 2011-2012 school year. We matched this student data to the *LA Times* data on published teacher value-added scores. The *LA Times* provided us with its value-added scores for the August 2010 release and the May 2011 update. The 2010 release includes score for teachers who had taught 60 students or more, while the 2011 release includes scores for almost all the teachers.[7] The *LA Times* also gave us its number-of-students variable used to decide whether a teacher would be published in the initial release to the public in 2010.

We supplemented these data by matching schools in our sample with school-level demographic information provided by the Common Core of Data from the National Center for Education Statistics (NCES). This information includes demographic shares by race/ethnicity and those receiving free or reduced-priced lunches. We matched 97 percent of teachers to a school.[8]

Table 1 shows the summary statistics for our sample of teachers in the 2010-2011 school year. We break the teachers into three groups: published (column 2), not published (column 3), and teachers who are within the chosen bandwidth for the regression discontinuity analysis (column 4). The teachers have a diverse racial background and are primarily female (69.3 percent). On average, unpublished teachers have spent less time teaching overall and within the district, which is not surprising given that the publication was based on the number of students taught. The regression discontinuity sample is more similar to the not published teachers than the published teachers, which suggests teachers are similar around the publication cutoff.

We use a regression discontinuity design to estimate the effects of the publication of teachers' scores. The *LA Times* published teachers who had taught 60 or more students from spring 2003 to spring 2009 and did not publish teachers below that threshold. This publication rule is sharp: 100 percent of teachers who had taught 60 or more students had

[7] Scores for teachers with fewer than 10 students were not published.

[8] Teachers we could not match were missing the California-District-School codes associated with their school, which allow us to match their school to the NCES code contained in the Common Core of Data.

their scores published, while no teachers under this cutoff had their scores published.

In order for the regression discontinuity approach to be valid, publication status must not be manipulable. The publication of teachers' value-added scores was an unexpected event and teachers had no knowledge of the 60-student cutoff; it is unlikely they could have influenced whether they would have been published. However, it is possible that teachers who knew their scores would be published might leave the district. In Figure 3 we follow McCrary (2008) and plot the distribution above and below the cutoff. The distribution exhibits no significant discontinuity, which is consistent with no manipulation of treatment status by teachers and no differential attrition below the cutoff. We supplement this test with a density test developed by Cattaneo et al. (2016) that aims to improve size and power over other density-based manipulation checks. The p-value for that test of manipulation is 0.75, which is further evidence of no manipulation. This result is unsurprising given that teacher ratings and the decision to publish them was controlled by the *LA Times*.

On average, teachers near the 60-student cutoff should be similar and vary only in that those above the cutoff were published while those below the cutoff were not published. In Figure 4, we plot the baseline covariate values above and below the 60-student cutoff. The figure illustrates that teacher quality is the same both above and below the cutoff. More formally, we test the differences in available covariates for teachers using the same regression discontinuity specification we describe below for measuring impacts on outcomes. Table 2 shows these differences for ten available covariates. While none of the differences in teachers' characteristics above and below the cutoff is statistically significant at conventional levels, there are quantitatively large differences in several of the demographic characteristics, such as gender and race. We can also use our specification to conduct a joint test of whether baseline covariates predict publication by using a weighted regression of publication on covariates interacted with our running variable with a triangular kernel described below. An F-test of whether the coefficients on these covariates are different from zero cannot be rejected (p-value=0.41). Nonetheless, the tests of demographic differences across the discontinuity are underpowered. Reassuringly however, characteristics predictive of teacher performance are not statistically different from each other, including teachers' value-added and teachers' experience, and we will conduct falsification tests for all of our analyses as a separate check of the validity of our design as well.

We begin by examining the effects of publication of value-added scores measured at the teacher level. We estimate the effects of publication on classroom composition (sorting), teacher attrition and whether teachers change schools. To measure the effects of publication, we estimate the following weighted, local-linear regression around the 60-student cutoff using a triangular kernel:

$$\begin{aligned} Outcome_i = \beta_0 + \beta_1 \cdot Pub_i + \beta_2 \cdot students_i + \beta_3 \cdot teacherVA_i + \beta_4 \cdot Pub_i \cdot students_i + \\ \beta_5 \cdot Pub_i \cdot teacherVA_i + \beta_6 \cdot students_i \cdot teacherVA_i + \beta_7 \cdot Pub_i \cdot students_i \cdot teacherVA_i + \\ X_{ij}\Delta + students_i \cdot X_{ij}\Gamma + \varepsilon_i \quad (1) \end{aligned}$$

where Pub_i is an indicator equal to 1 if the teacher was published and $students_i$ is the (centered) number of students a teacher had taught up to 2009 as constructed by the *LA Times* to determine publication. The coefficient of interest is β_1, which is the effect of publication on $Outcome_i$, such as lagged student test scores or class size. We also include teachers' tenure status and years of experience to aid precision. We allow the coefficients to vary flexibly across the cutoff by including their interactions with the running variable ($students_i$). To measure teachers' value-added ($teacherVA_i$), we use the 2003-2010 average value-added score, which was calculated by the *LA Times* for all teachers using data prior to publication. This measure provides baseline teacher value-added scores for both published and unpublished teachers, and we interact this term with the indicator for publication to examine heterogeneity in effects. (We also include an interaction term of this variable with the running variable.)[9] We use the bandwidth-selection process formulated by Calonico,

[9]The published scores in 2010, calculated using 2003-2009 data, are highly correlated with the published score in 2011, which are calculated using 2003-2010 data. The correlation coefficient between the average published score in 2010 and the average

Cattaneo and Titiunik (2014). Note that the optimal bandwidth differs by outcome, which implies that the sample size for each outcome in a regression will differ. Following Lee and Card (2008), we cluster standard errors at the level of the running variable to account for potential specification error induced by a discrete running variable.

4 Publication Effects

Student Sorting

In this section, we estimate the extent to which the publication of teachers' scores changed the composition of classrooms. Specifically, we examine whether high (low) performing students were more likely to be in the classrooms of high (low) rated teachers. Table 3 Panel A shows the results of the regression discontinuity estimation on classroom composition as measured by students' lagged test scores. Math score results are presented in columns (1) and (2) and English scores in columns (3) and (4).

The overall effect of score publication on student composition in the classroom in columns (1) and (3) is positive but not statistically significant. In columns (2) and (4) we add an interaction between publication and baseline teacher quality, and we find positive matching between students and teachers. Published high value-added teachers are in classrooms with higher performing students than unpublished high valued-added teachers. Specifically, teachers with a half standard deviation higher value-added score taught students with three tenths of a standard deviation higher lagged math and English scores.

We also examine whether students are more likely to sort toward high-performing teachers or away from low-performing teachers. In Figures 5 and 6 we plot the lagged class test scores for English and math for teachers whose combined rating is above the mean. In Figures 7 and 8 we plot test scores for teachers whose combined rating is below the mean. The sorting effect can be seen for the above-mean value-added teachers; the teachers above

published score in 2011 is 0.95, even around the 60-student cutoff.

the publication cutoff teach classrooms with higher incoming student test scores than those teachers below the cutoff. There is no discernible effect for below-mean value-added teachers. This evidence shows that the results are driven by students sorting toward high performing teachers.

In the remaining columns of Table 3 we examine the effect of publication on class size. Published teachers are in classes with 0.34 more students than unpublished teachers, on average, which is not statistically significant. This effect does not significantly vary by teachers' value-added, as seen in column (6) of Table 3, though the coefficient is positive.

In Panel B, we conduct falsification tests for the year prior to publication (the 2009-2010 school year). We should find no effect of publication for this year as no publication occurred. No statistically significant effects are found for publication on the various outcomes considered in Panel B. In particular, the effects of the interactions between publication and teachers' value-added are small and insignificant with respect to student sorting.

In Figures 9 and 10, we show how robust the coefficient estimates of publication interacted with teacher value-added are to changes in bandwidth. The estimated effects are significant at the 95 percent level for bandwidths, plus or minus five students, around the estimated optimal bandwidth.

In the 2010-2011 school year, the publication of teacher value-added scores represented new information for all parties involved—parents, teachers, and administrators. Each of these parties may have had priors about the quality of a teacher, but these scores were the first quantitative evidence that was made available to them. Once the information was revealed, each of these parties may have had incentive to use this information to match high-performing students with high-rated teachers. For instance, there are several common strategies for assigning students and teachers to classrooms. Schools may assign students randomly, group high-scoring students together (tracking) and rotate which teachers get the high-performing students, or schools may assign high-scoring students to high-quality teachers (tracking and matching). A recent paper by Hedvig Horvath (2015) found that in

North Carolina about a third of schools have random assignment, a third of schools track students but rotate assigned teachers, and the remaining third of schools track and match good students to good teachers. An analysis by Lefren and Jacob (2007) using data from a different state and school district suggests that principals assign students to classes in an effort to achieve balance in race, gender, and ability. That same study suggests that parents have a mechanism by which to request certain teachers and that almost a quarter of parents exercise this option. Thus, there are several potential channels for altering classroom assignments as a result of publication: Principals who use the matching strategy to assign students to teachers could perceive that they have a greater ability to match teachers and students as a result of ratings release; teachers with high published value-added scores might have more bargaining power within their school to request assignments to higher-performing students; and engaged parents may have used the published scores to directly lobby schools to place their children with high value-added teachers.

These effects may be especially salient when teachers within a grade level at a school exhibit a significant degree of variation. For example, a parent whose child was placed with a low-performing or unpublished teacher could learn from the *LA Times* website that a much better teacher was available and request the school to move his or her child into that classroom. We test this hypothesis empirically. We calculate the difference between a teacher's value-added score and the average value-added score of his or her immediate peers, where the peers are teachers at the same school and grade level.[10] We then add an interaction between this difference and the indicator variable for publication as follows:

$$Outcome_i = \gamma_0 + \gamma_1 \cdot Pub_i + \gamma_2 \cdot students_i + \gamma_3 \cdot distance_i + \gamma_4 \cdot Pub_i \cdot students_i +$$
$$\gamma_5 \cdot Pub_i \cdot distance_i + \gamma_6 \cdot students_i \cdot distance_i + \gamma_7 \cdot Pub_i \cdot students_i \cdot distance_i +$$

[10]Formally, we calculate the leave-out, mean value-added score by school and grade level for each teacher. We assign teachers with no peers a distance score of zero. Less than 3 percent of teachers have no peers.

$$X_{ij}\Delta + students_i \cdot X_{ij}\Gamma + \varepsilon_i \quad (2)$$

where $distance_i$ is the difference calculation referenced above for teacher i. As in the specification in equation 1 all variables are interacted with the running variable so that effects can vary flexibly across the cutoff. Table 4 Panel A shows the results of this estimation. The coefficient on the interaction between publication and the distance of a teacher to his or her immediate peers is positive and statistically significant, meaning that published teachers whose scores are much higher than those of other teachers at their school in the same grade level are more likely to have high-achieving students. The estimated coefficients are not qualitatively different from those found in Table 3. These results suggest that a teacher's score relative to his or her peers influences classroom composition. This result may be due to the timing of publication, which occurred late in the summer. This late timing may have restricted the ability of parents to shift their children across schools or for schools to change personnel, leaving within-school moves as the primary method for students to sort. Panel B of Table 4 shows the results of a falsification test similar to the one described for Table 3 Panel B where 2009-2010 is assumed to be the publication year instead of 2010-2011. No statistically significant results are found in Panel B, suggesting the results in Panel A are not driven by unobserved factors.

Teacher Attrition

As discussed in the introduction, the implications for teacher performance and attrition are ambiguous. Theory suggests there could be a positive effect of publication on teacher effort and performance driven by reputation concerns, intrinsic motivation or professional development. In contrast, negative effects could occur due to the crowding out of intrinsic incentives by extrinsic, reputational rewards or demoralization, or because highly rated teachers shift their attention from improving test scores to improving aspects of learning outside the scope of standardized tests. The effect of publication may be further heterogeneous as teachers adjust their effort after learning of their status relative to their peers. Moreover, teachers may wish or be compelled to switch schools, to switch out of tested grades, or to leave the district.

Using the regression discontinuity design above, we first study the effects of publication on whether teachers leave the district or tested grades or switch schools. We define a teacher as having left the sample if he or she is no longer found in the LAUSD administrative data in either 2011 or 2012 or in an untested grade level during those years, but was present in the baseline year of 2010. We define school switching similarly: A teacher has changed schools in either 2011 or 2012 relative to the school he or she taught at in 2010.

Table 5 presents the impact of publication on both attrition from the data (Panel A), school switching (Panel B), and moving to an untested grade level (Panel C). Columns (1) and (3) show the main effects of publication in 2011 and 2012, respectively. Columns (2) and (4) test for differential effects by teacher value added for 2011 and 2012 as well. In Panel A and Panel B published teachers are generally less likely to be retained. The effects in Panel C—moving to an untested grade—are essentially zero. We do see teachers are significantly less likely to be retained in 2011, but this effect dissipates by 2012. One interpretation is that publication causes teachers marginally attached to the profession to leave the district sooner than they may have otherwise. The interaction with teachers' value added is positive though not significant. Rockoff et al. (2012) found that, after informing principals about teachers' value added, low value-added teachers were 2-3 percentage points more likely to exit the school. Our confidence intervals in both 2011 and 2012 contain this effect size as well.

5 Conclusion

The *LA Times* published value-added ratings because, "research has repeatedly found that teachers are the single most important school-related factor in a child's education. Until now, parents have had no objective information about the effectiveness of their child's teacher."[11] The goal was to provide parents information about their children's teachers and allow families to respond to this information. In general, it can be difficult to introduce this type of information in the public sector. Unions are often opposed to these measures. Similar to several other contexts, the *LA Times* used a Freedom of Information Act and a contract with an economist to calculate value-added scores, which circumvented both the district and the teachers' unions.

We examine the effects of this publication using a regression discontinuity design. We find that students with higher test scores sort into classrooms with higher rated teachers. This sorting is responsive to significant ratings disparities across teachers within grades in the same school. We find some evidence of effects of publication on measures of teacher attrition.

It is unclear whether these results are desirable from a policy standpoint. One the one hand, both economists and policymakers want to ensure that families make informed decisions. As cited in our introduction, research has shown that families, including low-income families, have strong preferences for higher performing schools and teachers. On the other hand, there are several reasons for caution. First, putting information online may not foster equitable access to this information. For instance, recent work by Bergman (2017) found that the vast majority of families never access online information about their child's missed assignments and grades when available. Those who do access this information tend to have higher incomes and higher performing children. Second, there is already substantial concern around the ability to generate unbiased value-added estimates due to student sorting into classrooms (Rothstein, 2010). This problem may be exacerbated by publishing value-added scores if students begin to sort into classrooms based on more difficult-to-observe characteristics. While it is difficult to address the latter, we suggest that if school districts do want to publish teachers' value-added, they should actively push this information to all families

[11] See: `http://projects.latimes.com/value-added/faq/#publish_scores`

and ensure that everyone has an equal capacity to respond to this information accordingly.

References

[1] Adnot, Melinda, Thomas Dee, Veronica Katz, and James Wyckoff. "Teacher Turnover, Teacher Quality, and Student Achievement in DCPS." Educational Evaluation and Policy Analysis 39.1 (2017): 54-76.

[2] Bates, Michael. "Public and private learning in the market for teachers: Evidence from the adoption of value-added measures." University of California, Riverside Working Paper (2016).

[3] Bergman, Peter. "Nudging Technology Use: Descriptive and Experimental Evidence from School Information Systems." CESifo Working Paper Series No. 6101 (2016).

[4] Black, Sandra E. "Do better schools matter? Parental valuation of elementary education." Quarterly Journal of Economics (1999): 577-599.

[5] Buddin, Richard. "How effective are Los Angeles elementary teachers and schools?" (2010): 1-21.

[6] Buddin, Richard. "Measuring teacher and school effectiveness at improving student achievement in Los Angeles elementary schools." (2011).

[7] Buddin, Richard, and Gema Zamarro. "Teacher qualifications and student achievement in urban elementary schools." Journal of Urban Economics 66.2 (2009): 103-115.

[8] Calonico, Sebastian, Matias D. Cattaneo, and Roco Titiunik. "Robust data-driven inference in the regression-discontinuity design." Stata Journal 14.4 (2014): 909-946.

[9] Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. American Economic Review, 104.9 (2014): 2633-2679.

[10] Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. “Teacher-student matching and the assessment of teacher effectiveness.” Journal of Human Resources 41.4 (2006): 778-820.

[11] Cullen, Julie Berry, Cory Koedel, and Eric Parsons. “The Compositional Effect of Rigorous Teacher Evaluation on Workforce Quality. No. w22805. National Bureau of Economic Research (2017).

[12] Deming, David J. “Better schools, less crime?” The Quarterly Journal of Economics 126.4 (2011): 1-53.

[13] Figlio, David N., and Maurice E. Lucas. “What’s in a grade? School report cards and the housing market.” American Economic Review (2004): 591-604.

[14] Hanushek, Eric A. “The economic value of higher teacher quality.” Economics of Education Review 30.3 (2011): 466-479.

[15] Horvath, Hedvig. “Classroom Assignment Policies and Implications for Teacher Value-Added Estimation.” Draft. (2015).

[16] Hui, T. Keung. “Its teacher shopping season: Principals gently wield veto power over parents who request popular teachers.” News and Observer, July 15, 2013: p.1.

[17] Imberman, Scott A., and Michael F. Lovenheim. “Does the Market Value Value-Added? Evidence from Housing Prices After a Public Release of School and Teacher Value-Added.” Journal of Urban Economics 91 (2016): 104-121.

[18] Jackson, C. Kirabo. “Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina.” No. w18624. National Bureau of Economic Research (2012).

[19] Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger. “Teacher Effects and Teacher Related Policies. Annual Review of Economics 6 (2014): 801-825.

[20] Jacob, Brian A., and Lars Lefgren. "Can principals identify effective teachers? Evidence on subjective performance evaluation in education." Journal of Labor Economics 26.1 (2008): 101-136.

[21] Lee, Linda. "What Did the Teachers Think? Teachers' Responses to the Use of Value-Added Modeling as a Tool for Evaluating Teacher Effectiveness." Journal of Urban Learning, Teaching, and Research 7 (2011): 97-103.

[22] Lee, David S., and David Card. "Regression discontinuity inference with specification error." Journal of Econometrics 142.2 (2008): 655-674.

[23] Lefgren, Lars and Brian A. Jacob. "What do parents value in education? An empirical investigation of parents' revealed preferences for teachers." The Quarterly Journal of Economics 122.4 (2007): 1603-1637.

[24] McCrary, Justin. "Manipulation of the running variable in the regression discontinuity design: A density test." Journal of Econometrics 142.2 (2008): 698-714.

[25] Pope, Nolan. "The Effect of Teacher Ratings on Teacher Performance." Draft. 2014.

[26] Pop-Eleches, Cristian, and Miguel Urquiola. "Going to a Better School: Effects and Behavioral Responses." American Economic Review, 103(4): 1289-1324. 2013.

[27] Rockoff, Jonah E. "The impact of individual teachers on student achievement: Evidence from panel data." American Economic Review 94.2 (2004): 247-252.

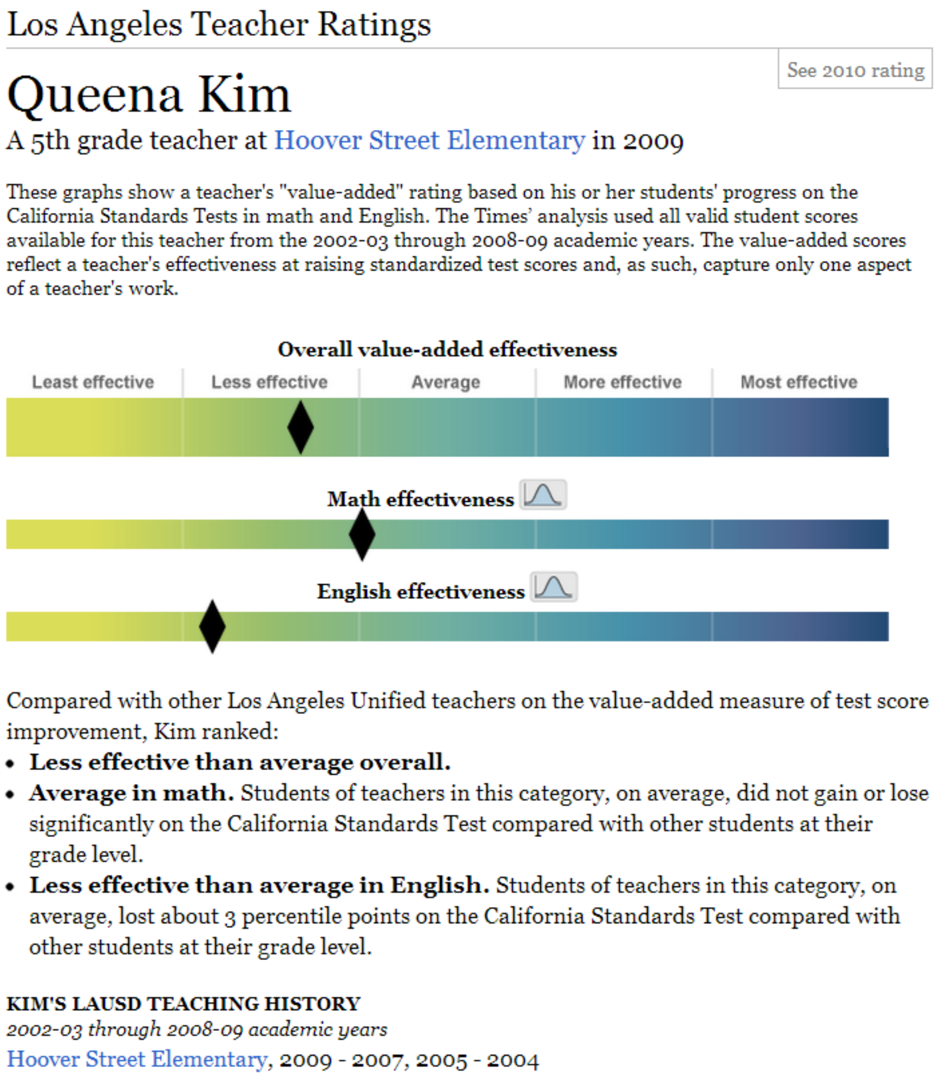
[28] Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools. American Economic Review, 102.7 (2012): 3184-3213.

[29] Rothstein, Jesse. "Teacher quality in educational production: Tracking, decay, and student achievement. Quarterly Journal of Economics, 125.1 (2010): 175-214.

[30] Song, Jason. "Teachers Blast L.A. Times for Releasing Effectiveness Rankings." *Los Angeles Times*. August, 30, 2010.

[31] Song, Jason and Jason Felch. "L.A. Unified releases school rating using 'value-added' method." *Los Angeles Times*. April, 12, 2011.

Figure 1: Example of an *LA Times* Teacher Page



Los Angeles Teacher Ratings

See 2010 rating

Queena Kim

A 5th grade teacher at Hoover Street Elementary in 2009

These graphs show a teacher's "value-added" rating based on his or her students' progress on the California Standards Tests in math and English. The Times' analysis used all valid student scores available for this teacher from the 2002-03 through 2008-09 academic years. The value-added scores reflect a teacher's effectiveness at raising standardized test scores and, as such, capture only one aspect of a teacher's work.

Compared with other Los Angeles Unified teachers on the value-added measure of test score improvement, Kim ranked:

- **Less effective than average overall.**
- **Average in math.** Students of teachers in this category, on average, did not gain or lose significantly on the California Standards Test compared with other students at their grade level.
- **Less effective than average in English.** Students of teachers in this category, on average, lost about 3 percentile points on the California Standards Test compared with other students at their grade level.

KIM'S LAUSD TEACHING HISTORY

2002-03 through 2008-09 academic years

Hoover Street Elementary, 2009 - 2007, 2005 - 2004

Notes: 2009 ranking from *LA Times* webpage

Figure 2: Teacher Value-Added Release Timeline

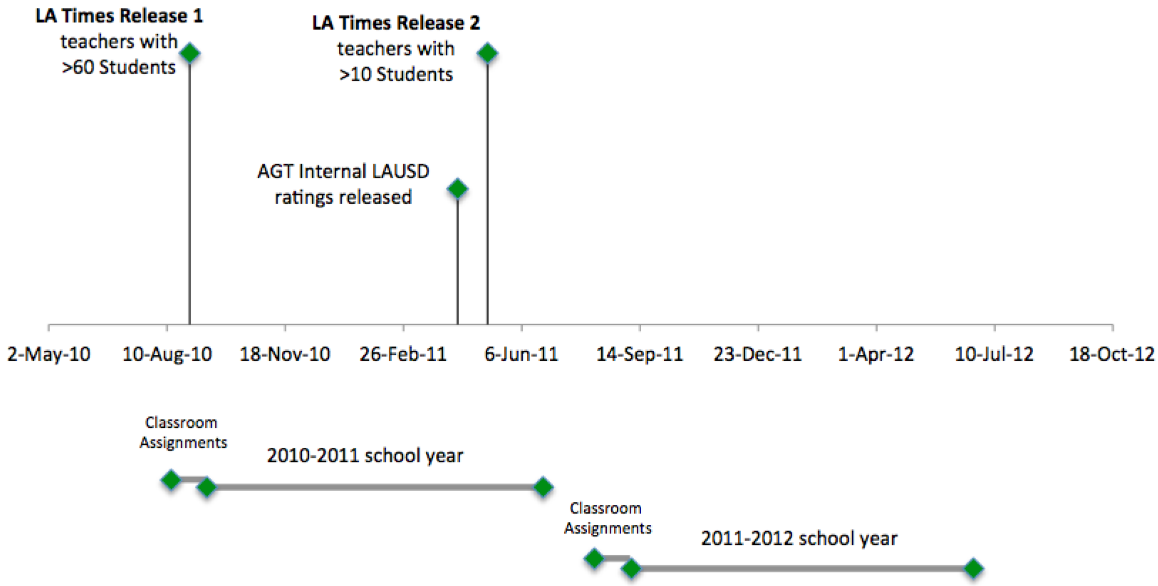
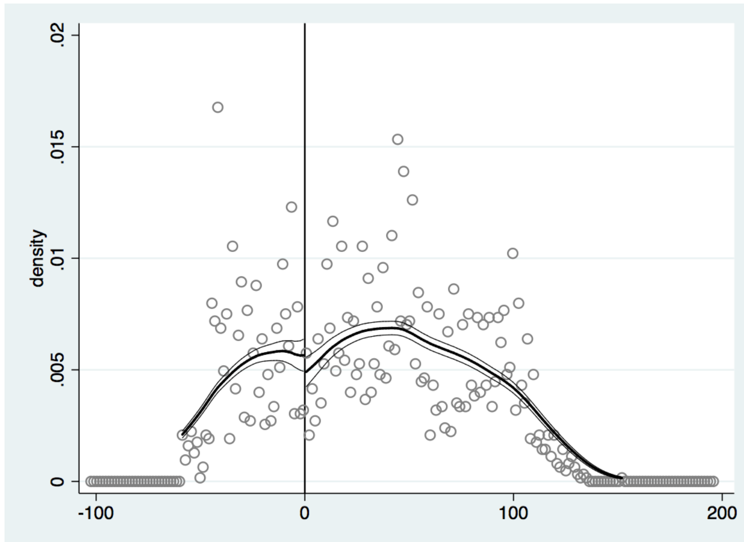
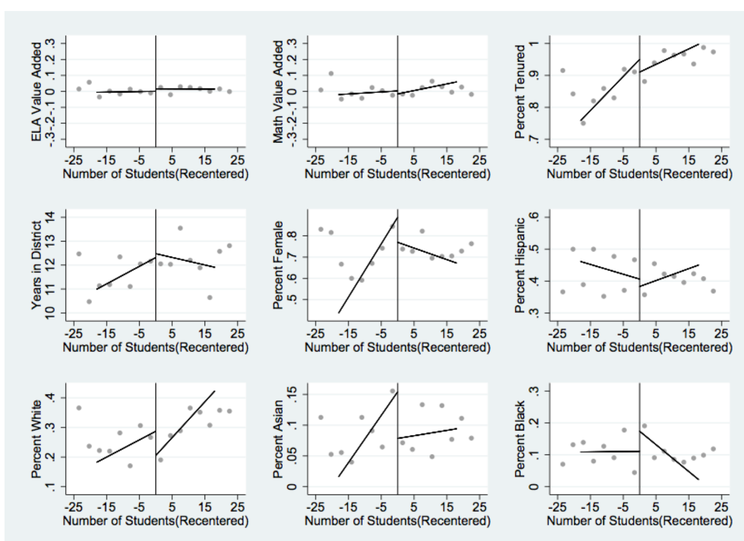


Figure 3: Density Test Around Cutoff



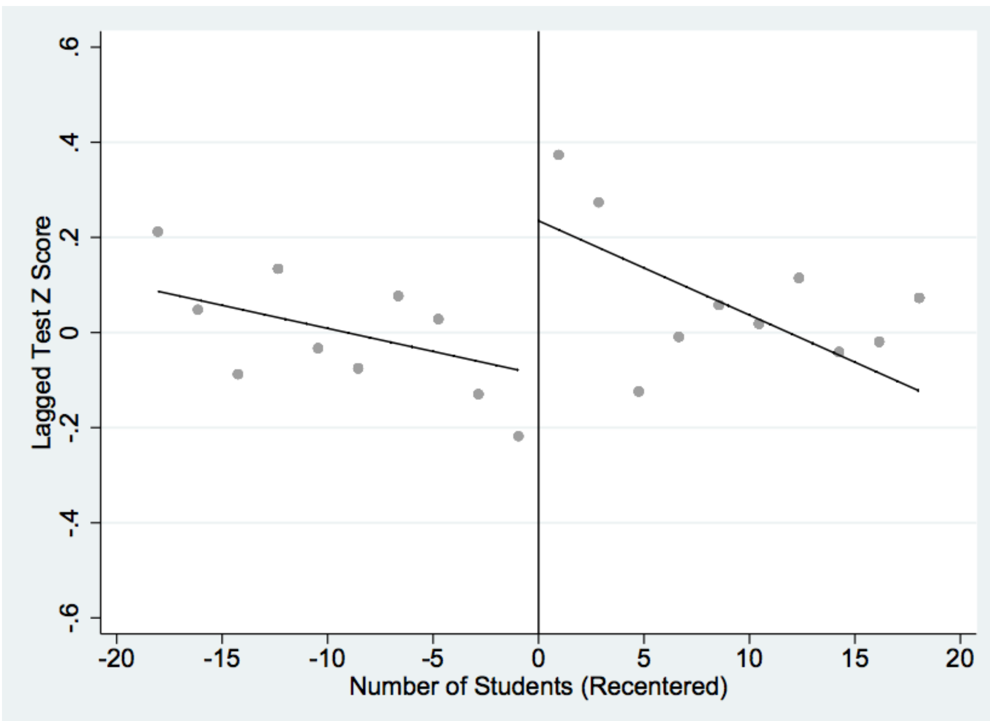
This figure shows the density test proposed by McCrary (2008). We plot the density of observations by the assignment variable (number of students taught with a test score and a lag score). We have re-centered the distribution so that 0 corresponds to the 60-student cutoff point used to determine publication. Data are from the Los Angeles Unified School District and the *Los Angeles Times*.

Figure 4: Baseline Covariates Above and Below the Cutoff



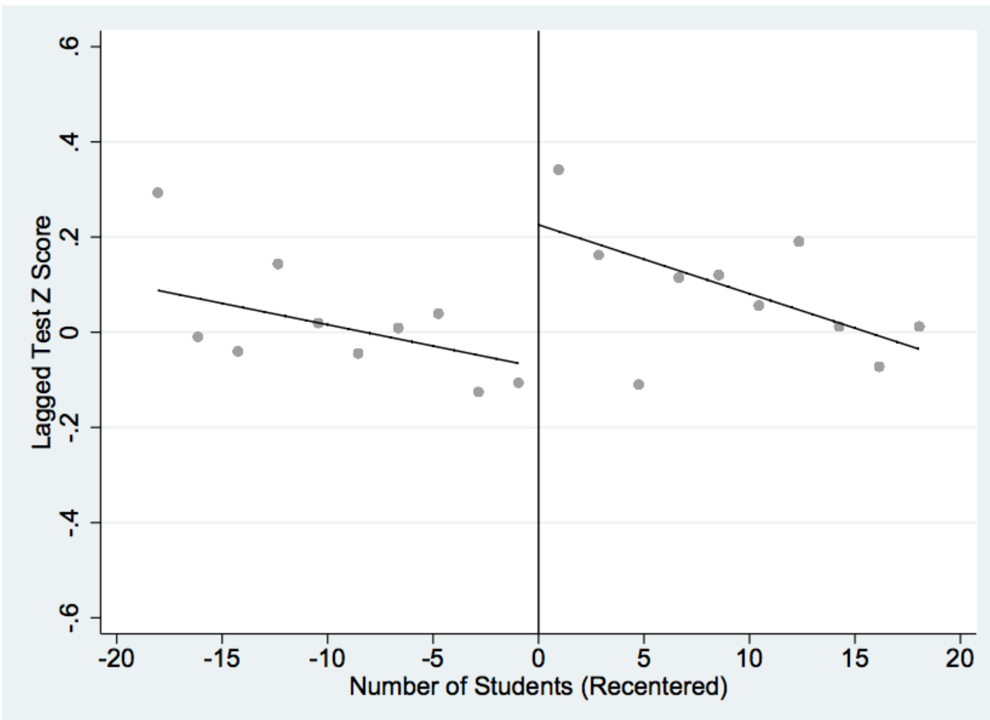
This figure plots baseline covariates, female, race/ethnicity indicators, math and English Language Arts (ELA) value-added, tenure status, and years in the district, above and below the 60 student cutoff. Baseline covariates are from the year 2010. Value-added scores and the number of students per teacher are from the *Los Angeles Times*.

Figure 5: Classroom Lagged Math Test Score for Above-Mean Rated Teachers



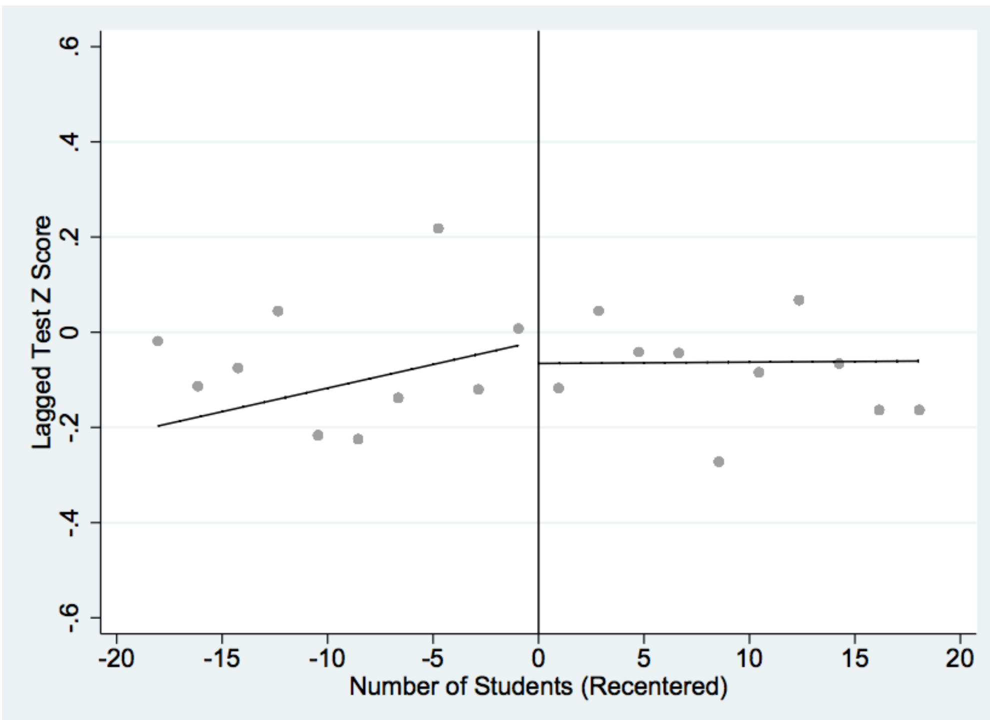
This figure plots students' lagged math test scores with respect to the re-centered running variable (the number of students calculated by the *LA Times*) and the associated regression lines using the specification described in the text. This graph is for teachers whose baseline value-added score is above the mean.

Figure 6: Classroom Lagged English Test Score for Above-Mean Rated Teachers



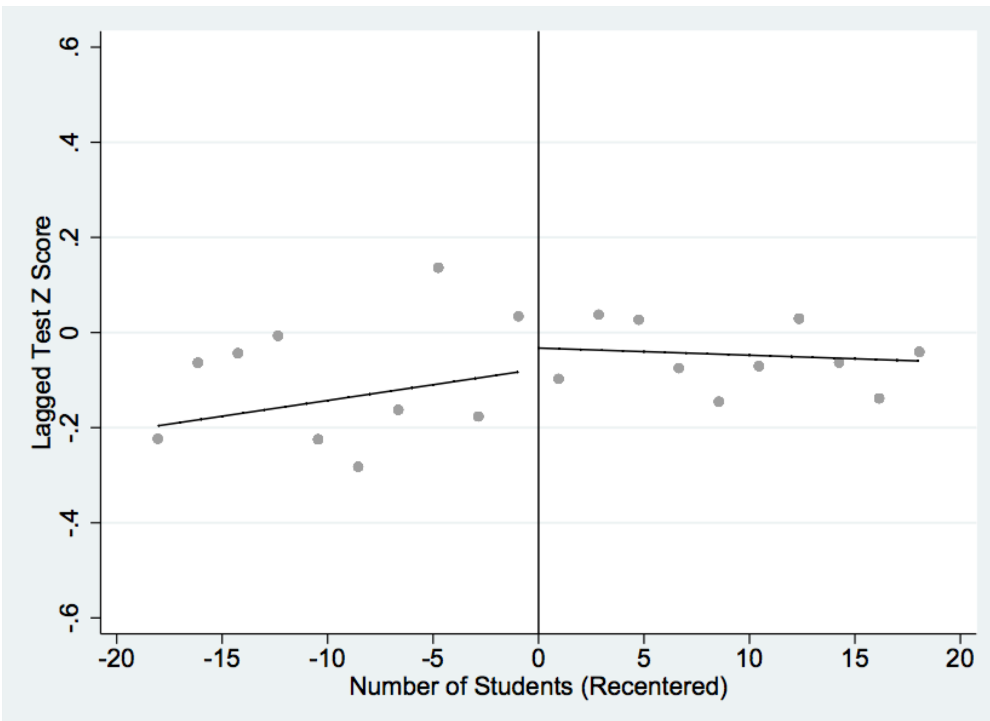
This figure plots students' lagged ELA test scores with respect to the re-centered running variable (the number of students calculated by the *LA Times*) and the associated regression lines using the specification described in the text. This graph is for teachers whose baseline value-added score is above the mean.

Figure 7: Classroom Lagged Math Test Score for Below-Mean Rated Teachers



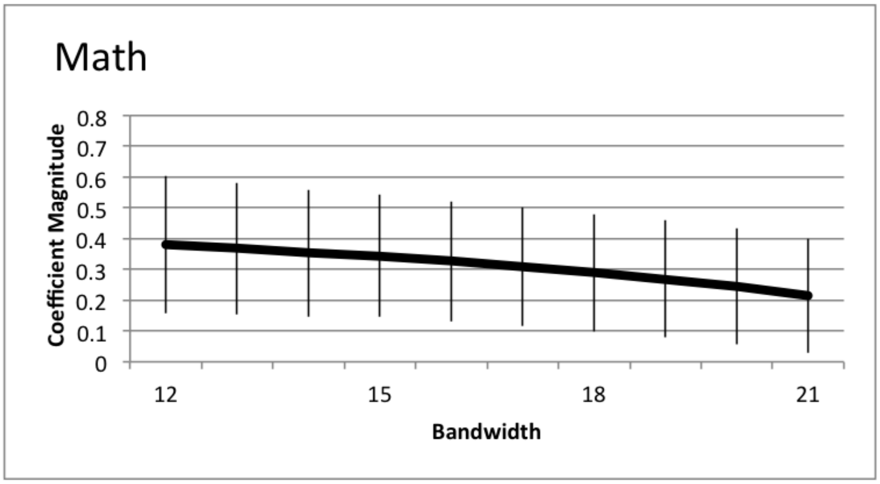
This figure plots students' lagged math test scores with respect to the re-centered running variable (the number of students calculated by the *LA Times*) and the associated regression lines using the specification described in the text. This graph is for teachers whose baseline value-added score is below the mean.

Figure 8: Classroom Lagged English Test Score for Below-Mean Rated Teachers



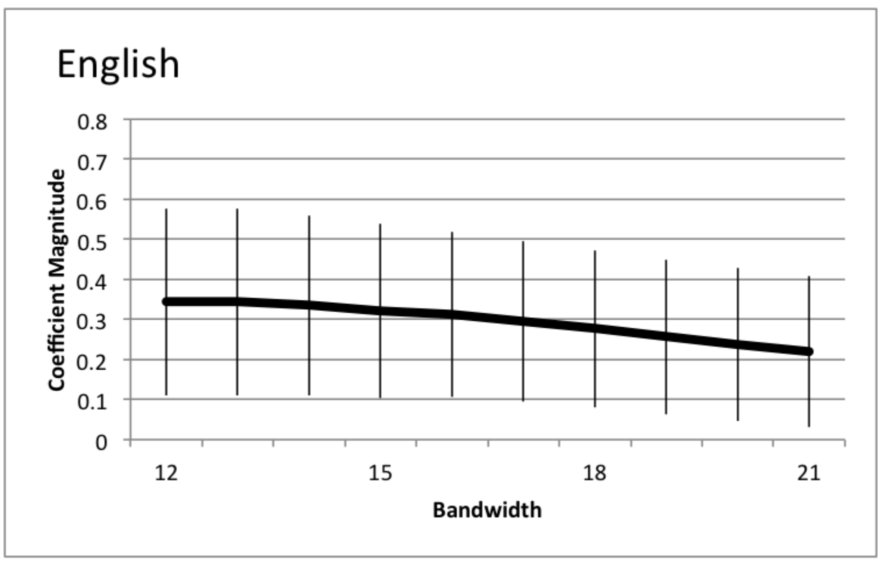
This figure plots students' lagged ELA test scores with respect to the re-centered running variable (the number of students calculated by the *LA Times*) and the associated regression lines using the specification described in the text. This graph is for teachers whose baseline value-added score is below the mean

Figure 9: Coefficient Estimates of Published Interacted with Value-Added for Math



This figure shows the coefficients on the indicator for published interacted with a teacher's value-added score. The dependent variable is the lagged math score of students in the teacher's classroom. Coefficients are shown for bandwidth +/- 5 students around the estimated optimal bandwidth described in the text. The specification is the one used in Table 3. Error bars depict 95 percent confidence intervals.

Figure 10: Coefficient Estimates of Published Interacted with Value-Added for English



This figure shows the coefficients on the indicator for published interacted with a teacher's value-added score. The dependent variable is the lagged English score of students in the teacher's classroom. Coefficients are shown for bandwidth +/- 5 students around the estimated optimal bandwidth described in the text. The specification is the one used in Table 3. Error bars depict 95 percent confidence interval.

Table 1: Summary Statistics

Teachers in 2011				
	All	Published	Not Published	RD Sample
Years in District	13.7	14.2	11.8	12.0
Years Teaching	14.1	14.7	12.1	12.2
White	37.6%	40.9%	26.9%	27.5%
Black	10.6%	10.9%	9.5%	10.4%
Asian	8.3%	8.5%	7.7%	8.7%
Hispanic	36.2%	34.0%	43.2%	42.8%
Female	69.3%	68.5%	71.9%	69.7%
Tenured	94.8%	97.6%	86.0%	90.8%
Published in 2011	75.9%	100%	0%	51.2%
Observations	3657	2777	880	676

Sample is teachers for whom the *LA Times* calculated value-added scores and who taught in the 2010-2011 school year. The regression discontinuity (RD) sample is teachers within the calculated bandwidth for the math score results.

Table 2: Teacher Covariate Balance

Variable	Difference	S.E.	p-value	Observations
Years in District	0.079	1.16	0.946	628
White	-0.096	0.085	0.260	676
Black	0.074	0.067	0.268	676
Asian	-0.077	0.064	0.226	676
Hispanic	-0.028	0.096	0.774	676
Female	-0.125	0.078	0.108	676
Tenured	-0.046	0.055	0.409	676
Value-Added in Math	-0.019	0.040	0.635	676
Value-Added in ELA	0.016	0.026	0.532	676

This table shows the balance of baseline teacher covariates across the publication threshold. The specification is the local-linear regression described in the text that is also used to assess impacts of publication on outcomes. The optimal bandwidth calculation is estimated as specified in the text. Robust standard errors shown.

Table 3: Student Sorting

Panel A: 2010-11						
	(1)	(2)	(3)	(4)	(5)	(6)
	Lagged Math Scores		Lagged English Scores		Class Size	
Publish	0.124	0.164	0.156	0.192*	0.337	0.353
	(0.112)	(0.107)	(0.097)	(0.097)	(0.499)	(0.509)
Publish*Teacher VA		0.316***		0.287***		0.435
		(0.099)		(0.101)		(0.582)
Observations	676	676	676	676	1,232	1,232
Panel B: 2009-10 (Placebo)						
	(1)	(2)	(3)	(4)	(5)	(6)
	Lagged Math Scores		Lagged English Scores		Class Size	
Publish	0.115	0.075	0.064	0.040	-0.460	0.404
	(0.090)	(0.087)	(0.104)	(0.098)	(0.571)	(0.569)
Publish*Teacher VA		-0.044		0.007		0.908
		(0.106)		(0.119)		(0.587)
Observations	738	738	692	692	1,082	1,082

This table shows the coefficient estimates from equation 1. The sample is teachers in the school year 2010-2011 for panel A and 2009-2010 in panel B within the bandwidth estimated according to Cattaneo et al. (2014) described in the text. Sample sizes differ by outcome because the optimal bandwidth is a function of the outcome variable. All regressions include baseline teacher value-added, experience and tenure status. Standard errors clustered by number of students in parentheses.

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

Table 4: Student Response: Distance to Peers

	(1) Lagged Math Scores	(2) Lagged English Scores
Panel A: 2010-11		
Publish	0.092	0.117
	(0.115)	(0.099)
Publish*Distance to Peers	0.281***	0.218***
	(0.051)	(0.072)
Observations	662	662
Panel B: 2009-10 (Placebo)	(1) Lagged Math Scores	(2) Lagged English Scores
Publish	0.036	-0.017
	(0.114)	(0.107)
Publish*Distance to Peers	-0.017	0.071
	(0.101)	(0.112)
Observations	720	673

This table shows the coefficient estimates from equation 1 with the published variable interacted with a measure of a teacher's value-added score difference from the school-grade-level mean value-added score. The sample is teachers in the school year 2010-2011 within the bandwidth estimated according to Cattaneo et al. (2014) described in the text. Sample sizes differ by outcome because the optimal bandwidth is a function of the outcome variable. All regressions include baseline teacher value-added, experience and tenure status. Robust standard errors in parentheses.

Table 5: Teacher Effects: Teacher Attrition and School Switching

	(1)	(2)	(3)	(4)
Panel A.	**Attrition from Data**			
	In 2011		In 2012	
Publish	-0.070**	-0.067*	-0.020	-0.015
	(0.033)	(0.036)	(0.039)	(0.043)
Publish*Teacher VA		0.025		0.008
		(0.045)		(0.064)
Observations	983	983	1,111	1,111
Panel B.	**Changed Schools**			
	Δ in 2011 relative to 2010		Δ in 2012 relative to 2010	
Publish	0.029	0.029	-0.029	-0.034
	(0.054)	(0.055)	(0.046)	(0.049)
Publish*Teacher VA		0.017		-0.041
		(0.045)		(0.074)
Observations	1,366	1,366	1,104	1,104
Panel C.	**No Test Scores**			
	In 2011		In 2012	
Publish	-0.007	-0.009	0.008	-0.000
	(0.070)	(0.065)	(0.058)	(0.056)
Publish*Teacher VA		0.019		-0.085
		(0.108)		(0.059)
Observations	1,567	1,173	1,327	1,327

Using 2009-2010 as the base year, this table shows attrition from the data, whether teachers switched schools, and whether they teach in an untested grade. Columns (1) and (2) show these outcomes from 2010 to 2011 and Columns (3) and (4) show these outcomes from 2010 to 2012. The sample is teachers in the school year 2009-2010 within the bandwidth estimated according to Cattaneo et al. (2014). Sample sizes differ by outcome because the optimal bandwidth is a function of the outcome variable. All regressions include controls for baseline value-added, years of experience and tenure status. Teacher value-added is from the *LA Times* 2003-2010 value-added scores. Robust standard errors clustered by students in parentheses.

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