Working Papers

www.cesifo.org/wp

The Effects of Making Performance Information Public: Evidence from Los Angeles Teachers and a Regression Discontinuity Design

Peter Bergman
Matthew J. Hill

CESifo Working Paper No. 5383
Category 5: Economics of Education
May 2015

ISSN 2364-1428

CESifo *Working Paper No. 5383*

The Effects of Making Performance Information Public: Evidence from Los Angeles Teachers and a Regression Discontinuity Design

Abstract

In theory, the publication of performance ratings may improve performance through reputation concerns and peer effects or impede performance by demoralizing employees. This paper uses school-district data and a regression discontinuity design to answer how consumers and employees respond to making performance information public. We find that high-performing students sorted into classrooms with highly-rated teachers as a result of publication. Teachers who were published do not perform better or worse than teachers who were not published on average. This average effect is due to the heterogeneous impact of publication; highly-rated teachers perform worse following publication while low-rated teachers perform better. On net, the gap between high and low-performing students closes slightly as a result.

JEL-Code: J240, I200.

Peter Bergman
Teachers College
Columbia University
USA - 10027 New York NY
bergman@tc.columbia.edu

Matthew J. Hill
UCLA & RAND
Los Angeles / CA / USA
matthill@g.ucla.edu

We thank Leah Platt Boustan, Richard Buddin, Dora Costa, Jonah Rockoff, Eric Taylor and Gema Zamarro for their detailed feedback, as well as seminar participants at AEFP, Toulouse and UCLA. All errors are our own.

1 Introduction

Performance information is publicly available in a variety of settings. For instance, consumers publicly evaluate restaurants, medical professionals and other service-industry members by posting ratings online.[1] Increasingly, and perhaps more controversially, performance information has been made public in the education sector. This context is of particular importance because researchers have quantified the significant social and economic value generated by high-quality teachers and schools (Rockoff 2004; Chetty et al. 2014; Deming 2011; Hanushek 2011; Jackson 2013). Many school districts, including the two largest districts in the country, New York City and Los Angeles, make school report cards available online, and *GreatSchools* has formulated and published ratings for more than 200,000 schools across the country.[2] In a number of districts, including districts in Florida, Cleveland, New York and Los Angeles, individual teacher performance information has also been made public.

The effect of publication on teacher performance is theoretically ambiguous. There are a number of potentially opposing effects. On the side of improving performance, information can improve the ability to perform one's job by providing feedback to principals and teachers. For example, teacher ratings may inform principals about which teachers need additional professional development. Rockoff et al. (2012) find that the productivity of schools (as measured by test scores) increases when principals are made aware of the value-added scores of their teachers. This effect is primarily due to the removal of low performing teachers. In a similar fashion, principals may threaten low performing teachers with job separation after the publication of performance information and thereby induce greater effort. Intrinsically motivated teachers may learn about their abilities and seek improvement (Kolstad, 2013). In addition, making performance information public can motivate employees through reputation or self-image concerns (Benabou and Tirole, 2006; Besley & Ghatak 2005; Jin & Leslie, 2003). Lastly, low rated teachers may seek guidance from high rated teachers and thereby improve

[1] Online ratings websites include *Yelp* and *Angie's list*.

[2] See www.greatschools.org.

their performance (Jackson & Bruegmann 2009).

On the side of reducing performance, releasing performance information potentially adds an extrinsic incentive for employees to perform well that may crowd out intrinsic motivation in a profession with employees often perceived to be altruistically motivated (Frey & Jegen 2001; Benabou & Tirole 2003). Intrinsic motivation may be important in a field that has few monetary rewards for improved performance (Jackson et al., 2014). High value-added teachers may be particularly susceptible to his crowd out if intrinsic motivation is positively correlated with value-added scores. If crowd out occurs, highly rated teachers may actually decrease their effort after learning about their quality relative to their peers. Moreover, publishing performance information may cause demoralizing stress that reduces effort (Lee, 2011).

Parents also may respond to the release of teachers' value-added scores. Previous research shows that parents react to teacher and school quality (Jacob & Lefgren, 2007; Black, 1999; Figlio & Lucas, 2004). Imberman and Lovenheim (2013) study whether housing markets responded to the publication of schools' value-added in Los Angeles. They find that neither publication is capitalized into housing prices. However, 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-quality teachers. Therefore information on teachers could encourage parents to "vote with their feet" and request higher-quality teachers. For instance, Jacob and Lefgren (2007) find that some parents will petition schools for their children to have higher-performing teachers. This shifting may occur differentially by student ability and potentially increase achievement disparities.

This paper examines the impact of publishing performance information on student sorting and teacher performance. In August of 2010, the Los Angeles Times (LA Times) published value-added scores for third through fifth grade teachers in Los Angeles Unified School District (LAUSD). The scores were computed for both English and math by an independent party contracted by the Los Angeles Times. Only teachers who had taught 60 students with test scores and lagged test scores up until spring 2009 were published.[3] In June 2011, the LA Times published scores based on data through spring 2010 for all third, fourth, and fifth grade teachers (almost) irrespective of the number of students the teacher had taught previously. The LA Times heralded the test scores with a front page article and provided the public with access to the scores via an online database.

Empirically, we examine the student and parent or school response to publication using a regression discontinuity design around the 60 student cutoff. We find that, in general, students with high test scores sort into classrooms with teachers that have high value-added scores. This finding is consistent with students and parents who value academic performance seeking out the higher-rated teachers, though schools could shift students across classrooms as well.

Student sorting and mean reversion complicate the analysis of the impact on teacher performance, however we find evidence of heterogeneous effects of publication on student test scores. We find that highly rated teachers appear to perform worse following publication while low-rated teachers perform better. Nolan Pope (2014) also looks at the publication of test scores in LAUSD on performance and finds similar effects for teachers.[4] These results are present for math scores in the first year of publication, but not for English. The strongest results are found in the second year of publication: the lowest rated teachers improve students' English and Math test scores by 0.05 standard deviations following publication, while the highest rated teachers' students' math and English scores decline by 0.07 to 0.04 standard deviations, respectively. We show that because these results occur two years after the value-added scores were calculated, they are likely the result of publication and not mean reversion. We do, however, demonstrate that teacher value-added scores exhibit significant mean reversion that dissipates quickly in our data.

[3]In order for a student to be usable in the LA Times value-added calculation, the student had to have a test score and a lagged test score. Therefore the 60 student cutoff means teachers must have taught 60 students with a test score and a lagged test score between spring 2003 and spring 2009.

[4]We became aware of Pope's work as we distributed our paper. A key difference between our paper and Pope is the ability to look at student sorting using a regression discontinuity design, which is feasible because of the data we obtained from the LA Times.

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. Sections 4 and 5 show the student and teacher effects of the publication of scores, respectively. Section 6 concludes.

2 Background

In August 2010, the LA Times newspaper published teacher value-added scores for third through fifth grade teachers in the Los Angeles Unified School District. The value-added scores were constructed by Dr. Richard Buddin, a senior economist at the RAND Corporation, hired by the LA Times. Details on Buddin's methodology can be found in his white paper on the subject (Buddin 2010). Buddin uses methods commonly found in the literature: linear regression with teacher fixed effects controlling for student covariates (Jackson et al. 2014).[5] The value-added scores were based on student test score data from 2002-2003 to 2008-2009 obtained by the LA times via a Public Records Act request. These value-added scores were calculated using data up until *two school years* prior to publication in the 2010-2011 school year.

The LA Times introduced the scores with a front-page story and then provided the public with an online database of teachers and their corresponding value-added scores.[6] 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. The evaluation of a teacher consists of an overall score as well as a math and English score. 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 this type of information as LAUSD had not previously computed these scores.

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

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

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 9 months following the initial release (Imberman & Lovenheim 2013). 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, itself, 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 teacher's 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 students from 2002-2003 to 2008-2009, where students needed to have at lease one year with a test score and a lagged test score. This 60 student cutoff provides a natural experiment. Teachers right below the cutoff should be similar to those just above the cutoff and therefore allow the use of a regression discontinuity design to estimate the impact of publication on various outcomes. In May 2011, the LA Times updated their value added scores to include the 2009-2010 school year (Buddin 2011). For the updated scores, 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.

LAUSD subsequently 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 LA Times and are thus similar to the published scores. However, the LA Times scores use data from 2002-2003 to 2009-2010 while the LAUSD internal scores used data only from 2007-2008 to 2009-2010. LAUSD denoted these scores as Academic Growth over Time (AGT) and provided them to teachers and principals in the spring of 2011. This implies any results using the spring 2012 test scores must be interpreted in conjunction with this information release as well as the release the year before by the LA Times.

3 Data

The data used in the paper comes from two sources: LAUSD and the LA Times. We filed a public records request with LAUSD to obtain similar data to that which the LA Times used to calculate the value-added scores: identifiable teacher names linked to deidentified student test scores. We have these data for LAUSD students from the 2008-2009 school year through the 2011-2012 school year. The LA times then provided us with their value-added scores for the August 2010 release and May 2011 update. The LA Times also gave us the number of students variable used to decide whether a teacher was published in the initial release to the public.

Table 1 shows the summary statistics for our sample of students in the 2010-2011 school year and their associated teachers. For students, the only demographic information available to us is parents' education, however parents' education data is measured with questionable accuracy. According to conversations with the district, LAUSD does not believe it is reliable and decided against using it in their internal value-added calculations due to quality concerns. We have more detailed demographic information for the teachers. The teachers have a diverse racial background and are primarily female (69.9 percent). 3,089 teachers who were published were still teaching in the 2010-2011 school year. There were 1,342 teachers for whom the LA Times calculated the value-added score but did not publish in 2010 because they fell below the 60 student cutoff.

4 Student Response to Publication

The LA Times publication was parents' first exposure to teacher value-added scores. There may have previously existed some local knowledge or perceptions about teacher quality, but the LA Times scores would have represented the first quantitative measurement of teacher quality available to parents. In this section we examine the student/parent response to the publication of the teacher value-added scores. We explore whether the data is consistent with parents lobbying to have their children placed in higher-performing teacher's classrooms and if this effect varies by student quality, though we cannot rule out whether school leadership re-sorts students without communication with parents.

We employ a regression discontinuity design to estimate the parental response to the publication of teacher scores. The LA Times published teachers who had 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 with 60 or more students were published while no teacher under the cutoff was published.

In order for the regression discontinuity approach to be valid, agents must not be able to manipulate their treatment status. The publication of teacher value added scores was an unexpected event and teachers had no knowledge of the 60 student cutoff, therefore, it is unlikely they could have influenced whether they would have been published. In figure 2 we follow McCrary (2008) and plot the distribution above and below the cutoff point. The distribution exhibits no significant discontinuity, which is consistent with no manipulation of treatment status by teachers.

Teachers near the cutoff should be equivalent and vary only in that those just above the cutoff were published while those below were not. In figure 3, we plot one important key covariate, teacher value added scores, 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 all available covariates of teachers within 19 students of the cutoff.

Table 2 shows that of the eight available covariates, the difference between teachers above and below the cut off are only statistically significant for one: a dummy variable indicating the teacher is white. The table shows that there are few differences in teacher characteristics above and below the cutoff, in particular teacher have similar experience, which is one of few teacher characteristics that correlates with value added (cf. Rockoff, 2004).

We estimate the following weighted, local-linear regression around the 60-student cutoff using a triangular kernel:

$$Outcome_i = \alpha_0 + \beta_0 \cdot Pub_i + \beta_1 \cdot students_i + \beta_2 \cdot Pub_i \cdot students_i + \varepsilon_i \qquad (1)$$

where Pub_i is an indicator equal to 1 if the teacher was published and $students_i$ is the 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_0, which is the effect of publication on $Outcome_i$.

Table 3 shows the results of the regression discontinuity estimation when a bandwidth of 19 students is used.[7] In columns (1)-(4) we investigate the effect of publication on student quality as measured by lagged-test scores (Math scores in columns (1) and (2) and English scores in columns (3) and (4)). No overall effect of publication on student quality is found (columns (1) and (3)). However, in columns (2) and (4) we add an interaction between publication and baseline teacher quality. To measure teacher quality, we use the 2003-2010 average value added score, which was calculated for all teachers using data prior to publication, in order to have a teacher value score for both published and unpublished teachers.[8] The coefficient on this interaction is positive and statistically significant, meaning that published high-value-added teachers have higher-quality students than unpublished high-valued-added teachers. Specifically, teachers with a one standard deviation higher value

[7] We use the suggested algorithm from Calonico, Cattaneo, and Titiunik (2014) to determine optimal bandwidth.

[8] The published score in 2010, calculated using 2003-2009 data, is highly correlated with the published score in 2011, which is calculated using 2003-2010 data. The correlation coefficient between the average published score in 2010 and the average published score in 2011 is 0.95–even around the cutoff. All of our results are robust to using our own calculated value-added scores using data prior to 2010, which correlate similarly to the 2010 score.

added score had students with a 0.26 average standard deviation increase in lagged math scores and a 0.18 average standard deviation increase in lagged English scores. In columns (5) and (6) we examine the effect of publication on class size. Published teachers have larger classes, however these effects are not quite statistically significant.

Our findings on the effect of publication on student sorting are robust to changes in specification and pass a placebo test. Panel B of Table 3 shows the results of the same regressions in Panel A in the year before publication (the 2009-2010 school year). If the regression discontinuity design is valid, we should find no effect of publication for this year as no publication occurred. No statistically significant effect is found for publication or the interaction between publication and teacher value added. In Figure 4, we show how the coefficient estimates of published interacted with teacher value added change by bandwidth. With a few exceptions, the estimated effects are significant at the 90 percent level for bandwidths from 5 to 25 students.

The results suggest that high performing students were sorted into classrooms with known (i.e. published) high-value-added teachers after publication. The precise mechanism for this sorting is unknown. One possibility is that schools allocated high performing students into high value-added teachers' classrooms. Another possibility is that parents of high-achieving students lobbied schools to have their children placed with high-value-added teachers. This effect would be especially salient when teachers within a grade level at a school exhibit a significant degree of variation. In this case, 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 published score and the average published score of their immediate peers, where their peers are teachers at the same school and grade level.[9] We then add an interaction between this difference and the indicator variable for publication (we also fully interact it with all other variables in

[9] We assign a value of 0 to unpublished teachers. We also assign teachers with no peers a zero. The number of teachers with no immediate peers is relatively small, less than 3 percent.

the regression to maintain the flexible form). Table 4 shows the results of this estimation. The coefficient on the interaction between publication and the distance of a teacher to their immediate peers is positive and statistically significant, meaning that published teachers whose score is much higher than other teachers at their school in the same grade level are more likely to have high achieving students. The coefficient on the interaction between publication and teacher quality becomes statistically insignificant with the addition of this new interaction, which suggests that the observed effect found in Table 3 operates through a teacher's score relative to their immediate peers and not the shifting of students across schools in accordance with a teacher's score relative to the entire distribution.

5 Teacher Response

5.1 Measures of Attrition

As discussed in the introduction, the implications for teacher performance are ambiguous. Theory suggests there could be a positive effect of publication on teacher effort and performance driven by reputation concerns, intrinsic motivation or peer or school investments; in contrast, negative effects could occur due to the crowd 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 or leave the district.

Using the regression discontinuity design, we first study the effects of publication on whether teachers leave the district or switch schools. We define a teacher as having left the district if he or she is no longer found in the LAUSD administrative data in either 2011 or 2012 but were present in the baseline year of 2010. We define school switching similarly: a teacher changes schools in either 2010 or 2011 relative to the school they taught at in 2010.

Table 5 presents the impact of publication on both attrition from the data (Panel A) and school switching (Panel B). Columns (1) and (3) shows the main effect of publication in years 2011 and 2012, respectively. Columns (2) and (4) test for differential effects by teacher value added for 2011 and 2012 as well. In both Panel A and Panel B, all estimates are not significant. In general, the standard errors are large relative to the point estimates, so it is difficult to make conclusions about the potential size of the effects.

5.2 Test Scores

The regression discontinuity approach used to gauge student response is not valid for analyzing impacts on teacher performance. Regression discontinuity requires that baseline covariates be continuous around through cutoff. In the previous section we showed that this is not the case. High-value-added teachers above the publication cutoff had higher-performing students after publication. Therefore we estimate the effect of publication using difference-in-differences. The estimated effect will then be the overall effect of publication.

We estimate the following equation with our sample of LAUSD students and teachers:

$$TestScore_{ij} = \alpha + \beta_1 Pub_j + \beta_2 Year2011_i + \beta_3 Pub_j \cdot Year2011_i + \gamma lagTestScore_i + X_j'\delta + \varepsilon_{ij} \tag{2}$$

where $TestScore_{ij}$ is the test score of student i in teacher j's class. Pub_j is an indicator variable equal to one if teacher j is published and will account for any time invariant differences between published and non-published teachers. $Year2011_i$ is a 2011 school year fixed effect and will control for the overall effect of the 2010-2011 school year on test scores. The variable of interest will then be the interaction between the published indicator variable and the 2011 school-year indicator variable. We also include variables to account for teacher experience (years in district fixed effects), student characteristics (grade level fixed effects), year fixed effects and teacher fixed effects.[10] Standard errors are clustered at the teacher

[10]We also used an alternative specification where we controlled for teacher experience with a quadratic in the number of students a teacher had taught, and the results were unchanged.

level.

Table 6 panel A shows the results of estimating equation 2. Columns (1) and (2) show the results of publication on math scores, while columns (3) and (4) show the results for English scores. In columns (1)and (3) we limit the sample to students with a balanced panel of teachers in all three school years that our data covers (2009, 2010, 2011), while in columns (2) and (4), we do not. The results suggest that publication had no overall effect on teacher performance. On average, published teachers do no better or worse than unpublished teachers after publication, which suggests publication has no effect on performance or the effect of publication is heterogeneous. An alternative measure of treatment is to use the share published within a student's grade level, which helps understand the net effects of publication overall. Panel B reports the results when share published is used as the explanatory variable, again, no significant effects of publication are found.

In Table 7 we examine if the effect of publication differs by teacher quality. In Panel A, columns (1) and (3) include an interaction between the publication in 2011 and the value-added score. The coefficient on this interaction is negative and statistically significant for student math scores and negative but not statistically significant at conventional levels for student English scores. The negative coefficient suggests that highly-rated teachers perform worse following publication while low-rated teachers performed better. We further explore this relationship in columns (2) and (4) by including dummies for the quintile rank of the teacher's value-added score. These quintiles correspond to the LA Times categorization of teachers online (least effective, less effective, average, more effective, and most effective). In column (2) the coefficient on most effective is negative and statistically significant, which suggests that the highest-performing teachers drive the negative coefficient in column (1). On average, a top rated teacher lowers student test score by 0.06 standard deviations in the following year.

In Panel B we include the 2011-2012 school year. This school year also included two more publication effects: principals reported value-added scores to teachers and the LA

Times published value-added scores for all teachers instead of only teachers with more than 60 students. The variable of interest in this specification is, again, the teacher's published value-added score interacted with the year of publication. For the 2009-2010 school year, this variable will be zero because no one was published. For 2010-2011, this variable will be zero for the non-published, and the published valued-added for the published teachers. For 2011-2012, this variable will be published value-added score. The estimated coefficients for both English and math are larger when 2012 is included. In columns 2 and 4 we observe significant effects and both the top and bottom of the score distribution for both English and math. Bottom rated teachers increase student test scores by 0.046 and 0.052 standard deviations for math and English, respectively. Top rated teachers lower test scores by 0.07 and 0.04 standard deviations for math and English, respectively.

5.3 Persistence

As shown above, we found that there are heterogeneous effects of publication on teacher performance, do these effects persist after the student moves to the next grade? We examine the effects of publication on subsequent year's test scores by re-estimating equation 2 with test score in $t + 1$ as the dependent variable. Table 9 shows the results of this estimation, columns (1) and (2) are analogous to columns (1) and (3) of table 7, panel A, except the dependent variable is now test score in 2012 instead of 2011. The heterogeneous effects found in table 7 are not found for test scores in the subsequent year. This suggests that while publication effects student performance for students' whose teacher was published in that year, this effect decays and does not persist for students in the subsequent year.

5.4 The High-Performing v. Low-Performing Student Achievement Gap

The heterogeneous effects found for teachers could possibly have implications for the achievement gap between high and low performing students. We have shown there are two opposing effects on the achievement gap between high and low-performing students: first, relative to low-performing students, high-performing students sort into classrooms with highly-rated teachers, which could result larger test-score disparities; second, low-rated teachers improve performance while highly-rated teachers decrease performance (on average), which could narrow test-score disparities. We examine the net effect by interacting the share published with lagged student scores. Table 10 shows the results of adding this interaction to equation 2. The coefficient on this interaction is negative and significant, which implies that students with test scores above the mean are negatively impacted by publication while students below mean benefit from having a higher share of teachers published at their grade level. The results suggest that publication of all teachers in a grade level would increase a student whose test scores are one standard deviation by 0.04 standard deviations. Thus, publication of teacher value added scores, overall, reduces the achievement gap between low and high performing students slightly.

6 Robustness

6.1 Mean Reversion

Recall that we found, in general, highly-rated teachers perform worse following publication and there is some evidence that low-rated teachers improve their performance. However, this observed effect is also consistent with teacher performance following a mean-reverting process. In this section, we examine the robustness of the previously estimated effects.

We show that if the ratings were calculated using test scores immediately prior to the year of publication, the impacts of publication one year later would be biased by mean reversion. However, ratings were calculated by the LA Times using scores *two* years prior to publication. We show that mean reversion dissipates quickly and does not persist two years after ratings are calculated.

We begin by performing a placebo test for the 2009-2010 school year—one year prior to publication—which demonstrates the presence of at least short-run mean reversion. In Table

8 we show the results of this exercise. The variable of interest is the triple interaction between an indicator variable equal to 1 if the teacher was published, a 2009-2010 year indicator variable and the published value added score. The specification in columns (1)-(4) is then equivalent to columns (1)-(4) in Table 7 except we treat 2009-2010 as the publication year instead of the actual publication year, 2010-2011. We find statistically significant coefficients of the triple interaction on both English and math scores. Columns (2) and (4) show that the tails of the distribution drive these results; low-rated teachers perform better while highly-rated teachers perform worse. The coefficients are the same sign as those found in the actual publication year and are larger. These significant "effects" are consistent with mean reversion, and since there was no publication in the year 2009-2010, it is likely mean reversion is driving the "effects" in columns (1)-(4).

In columns (5) and (6) we confirm that the results found in columns (1)-(4) are due to mean reversion. In columns (5) and (6) we replace the value added score (calculated with 2003-2009 data) in the triple interaction with the value added score the LA Times calculated using the years 2003-2010. This value added score includes the performance of the teacher in the 2009-2010 school year. We suspect that the results in column (1)-(4) stem from teachers with high/low realizations of value added reverting to their mean in 2009-2010. We then should see no effect when the score from the years 2003 to 2010 is used because it accounts for that reversion by including the teacher's 2009-2010 performance. We do not find any effect, which further suggests that mean reversion is indeed driving the results of columns (1)-(4).

The results of this placebo test show that teacher value added is subject to mean reversion, the important question is whether mean reversion is still present *two* years after calculation. The value-added scores were calculated from 2003-2009 and their publication is in the 2011 school year. If teacher value added is a slow mean reverting process, then the effects of mean reversion may still be present two years after calculation. Teachers with a high realization in 2009 (and therefore a high value added score for the period 2003-2009) will then have a lower score in 2010 and could then again in 2011. If teacher value added were such a process this would explain both the findings in Table 8 and Table 7. We examine if this is the case empirically below.

6.2 Estimating Mean Reversion

We have established that teacher value added scores follow a mean-reverting process. In this section we try to understand the extent to which mean reversion contaminates the estimated effect of publication. The principle question is the speed of mean reversion because the estimated effects are found two years after the calculation of teacher value added. If the speed of mean reversion is fast, any shock should dissipate after one year. Conversely, if the speed of mean reversion is slow, then we still may observe effects in the 2011 school year due to mean reversion. In order to estimate the speed of mean reversion we assume that teacher value added is an Ornstein-Uhlenbeck process as follows:

$$VA_t = VA_{t-1} + \lambda(\mu - VA_{t-1}) + \epsilon_t$$

Where VA_t is a teacher's value added score in a given year, t. VA_{t-1} is the score in the previous year and μ is the long-run average of the teacher's value added. The speed of mean reversion is then λ; as λ approaches 1, the speed of mean reversion is quicker and the previous realization exerts less influence on the present realization. We obtained student test score data in LAUSD from the years 2003-2011. With these data we calculated annual value added scores for teachers. Then we can estimate the speed of mean reversion parameter with the following regression:

$$S_{jt} = \lambda\mu_j + (1 - \lambda)S_{j,t-1} + \epsilon_{jt}$$

Where we approximate μ_i with the average of a teacher i's annual value-added scores across the years available. Note we are assuming λ is constant across teachers although each teacher will have a different value added, μ_j. With our 2003-2011 data, the estimated λ for math scores is 1.04 (standard error of 0.007) and 1.05 (standard error of 0.007) for English. This implies very fast mean reversion. After a positive shock, scores will have a small negative shock in the following period and no meaningful influence in subsequent periods.

The results of this estimation of the mean reversion parameter suggest that shocks to value added should only exert influence on the subsequent year's scores and no further. We can examine whether this is the case empirically. First we use our student test score data to calculate each teacher's value added score from 2003-2007, 2003-2008, and 2003-2009 (the 2003-2009 score will be equivalent to the published score). Then we can estimate the effect of these scores on teacher performance for one year following the calculation and for two years following the calculation, which will show the extent of mean reversion one year after calculated scores and two years after calculated scores respectively. Specifically, we use the following specification over the years t and $t-1$ to estimate the value added scores' impact:

$$TestScore_{ij} = \alpha + \beta * year_t * VA_j + \gamma lagTestScore_i + X_j'\delta + \epsilon_{ijt} \tag{3}$$

Where $TestScore_{ijt}$ is the test score of student i with teacher j. β is the coefficient of interest and will measure the impact of a teacher's value added score from 2003-year prior/two years prior on the present year's test scores. We include teacher, grade level and year fixed effects, as well as teacher experience controls.We estimate the above equation for t equal to 2011, 2010, and 2009. Figure 5 Panel A shows the estimates of β for math and Panel B shows the estimates for English. The black line shows the estimates when value added scores calculated up to one-year prior are used. For instance, in 2009, the value-added score from 2003-2008 is used. Effects in 2011 will be conflated with the publication effect. The figure shows that significant mean reversion exists. Value added calculated up to the prior year has a consistent statistically significant negative effect of around -0.03 for the non-publication years of 2009 and 2010. The effect is stronger in the publication year of 2011 likely because the estimate includes both the effect of publication and mean reversion. The dashed-gray line shows the effects of value-added scores calculated from 2003 to two-years prior. For instance, the value-added score from 2003-2007 is used in 2009. Given that we include teacher fixed effects, these value added scores should have no effect on test scores two years later if mean reversion dissipates after one year. For math, we only find a significant effect of prior value added in 2011, which is the publication effect found in Table 7 column (1). For English, similar to the estimation in Table 7, no effect is found in the publication year. A small, statistically significant effect is found in 2010 and none in 2009. The results suggest while mean reversion exerts a significant effect on performance one year after calculation, this effect dissipates by two years. Therefore the effect of publication on math scores found in Table 7 is likely due to publication and not mean reversion.

6.3 Top Coding

A possible reason for the diminished performance of high rated teachers following publication is the increased presence of top coded students in their class. Good students sort into highly-rated teachers' classrooms following publication. Some of these students may have had perfect scores on their previous year's tests. If this is the case, teachers will be unable to increase their students test scores because their is no margin for improvement. This phenomenon may be driving the negative effects found for high rated teachers. We examine this in table 11. In columns (1)-(4), we re-estimate equation 2 using a dummy equal to one if the student's lagged test score was top coded as the dependent variable. The coefficient on the interaction of the publication dummy and the teacher's value added score is positive and statistically significant for math in both the year 2011 and when the year 2012 is added to the estimation. No significant effects is found for English, likely because a perfect test score on English is a rare event: 2 percent of students achieve the highest score on math, while only 0.2 percent of students do so for English. The positive coefficient suggests that high;y-rated teachers were more likely to have top coded students. This result is expected given that we have shown good students sort into high rated teacher's classrooms. In columns (5)-(8)

we remove these top coded students from the sample and re-estimate our specification from table 7 to ascertain the extent to which the top coded students are influencing our results. The differences in the coefficient estimates in table 11 and their analogous estimates in table 7 are not statistically significant which suggests that the increased presence of top coded students is not driving the negative results for high rated teachers.

7 Conclusion

The publication of performance measures occurs in many fields. One set of theories imply that publication could incentivize performance improvement through reputational and self-image concerns, as well as the ability to learn about one's effectiveness and learn from others' effectiveness. However, publishing performance information can also have deleterious effects by angering and embarrassing employees or crowding out intrinsic motivation. Which of these theories dominates is an important empirical question. Moreover, it can be difficult to introduce performance incentives in public education. Unions are often opposed to these measures and the costs of providing incentives can be untenable in the face of budget constraints. In the example we study, 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 teachers' unions.

We find evidence for both positive and negative impacts. On average, there is no overall impact of publication on test scores, which is due to heterogeneous impacts on high versus low-rated teachers. The latter exhibit performance improvement while the former reduce performance. Among several possibilities, the reduction in performance could be due to the crowding out of intrinsic motivation or possible shifting effort away from test preparation. Key for the analysis of publication effects is the possible presence of mean reversion. We document that teacher value added scores exhibit significant mean reversion. Mean reversion complicates the evaluation of teachers and perhaps other public officials whose performance is rated. However, we show that this mean reverting process is not persistent—dissipating after one year.

Lastly, and perhaps most importantly, we find students with higher test scores sort into classrooms with higher-rated teachers. This sorting is particularly strong in grades with significant rating disparities across teachers within grades in the same school. If highly-rated teachers also have large impacts on higher-performing students, this sorting could increase achievement disparities. However, the overall net effect of these results is to diminish, albeit marginally, the achievement disparity between high and low-performing students, as measured by test scores.

References

[1] Benabou, Roland, and Jean Tirole. "Intrinsic and extrinsic motivation." The Review of Economic Studies 70.3 (2003): 489-520.

[2] Besley, Timothy, and Maitreesh Ghatak. "Status incentives." The American Economic Review (2008): 206-211.

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

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

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

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

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

[8] 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.

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

[10] Downes, Thomas A., and Jeffrey E. Zabel."The impact of school characteristics on house prices: Chicago 19871991." Journal of Urban Economics 52.1 (2002): 1-25.

[11] Dranove, David, and Ginger Zhe Jin. "Quality disclosure and certification: Theory and practice. No. w15644. National Bureau of Economic Research, 2010.

[12] 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.

[13] Frey, Bruno S., and Reto Jegen."Motivation crowding theory." Journal of Economic Surveys 15.5 (2001): 589-611.

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

[15] 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." No. w19157. National Bureau of Economic Research, 2013.

[16] 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.

[17] Jackson, C. Kirabo, and Elias Bruegmann. "Teaching students and teaching each other: The importance of peer learning for teachers. American Economic Journal: Applied Economics 1.4 (2009): 85-108.

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

[19] 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.

[20] Jin, Ginger Zhe, and Phillip Leslie."The effect of information on product quality: Evidence from restaurant hygiene grade cards." The Quarterly Journal of Economics (2003): 409-451.

[21] Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger. “School quality, neighborhoods, and housing prices.” American Law and Economics Review 8.2 (2006): 183-212.

[22] Kane, Thomas J., and Douglas O. Staiger. “Estimating teacher impacts on student achievement: An experimental evaluation.” No. w14607. National Bureau of Economic Research, 2008.

[23] Pope, Nolan. “The Effect of Teacher Ratings on Teacher Performance.” Draft. 2014.

[24] Kolstad, Jonathan T. ”Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards.” American Economic Review, (2013): 2875-2910.

[25] 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.

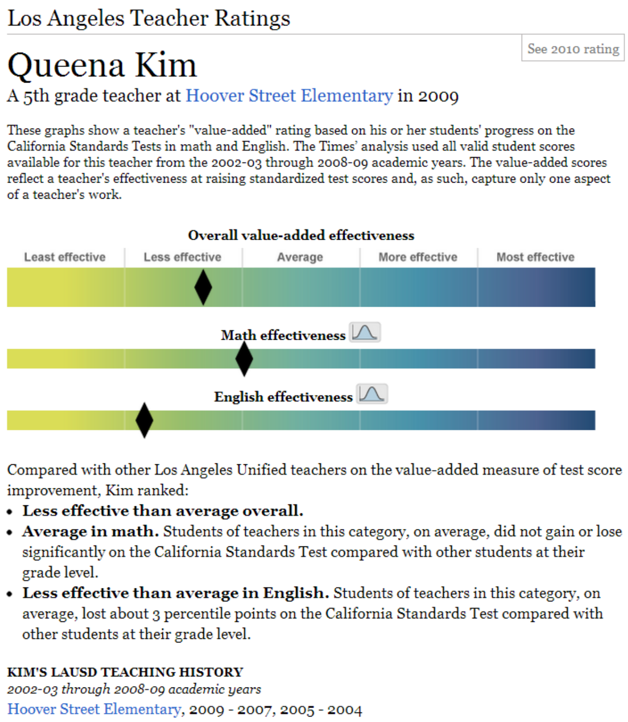
[26] McCrary, Justin. “Manipulation of the running variable in the regression discontinuity design: A density test.” Journal of Econometrics 142.2 (2008): 698-714.

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

Figure 1: Example of 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.

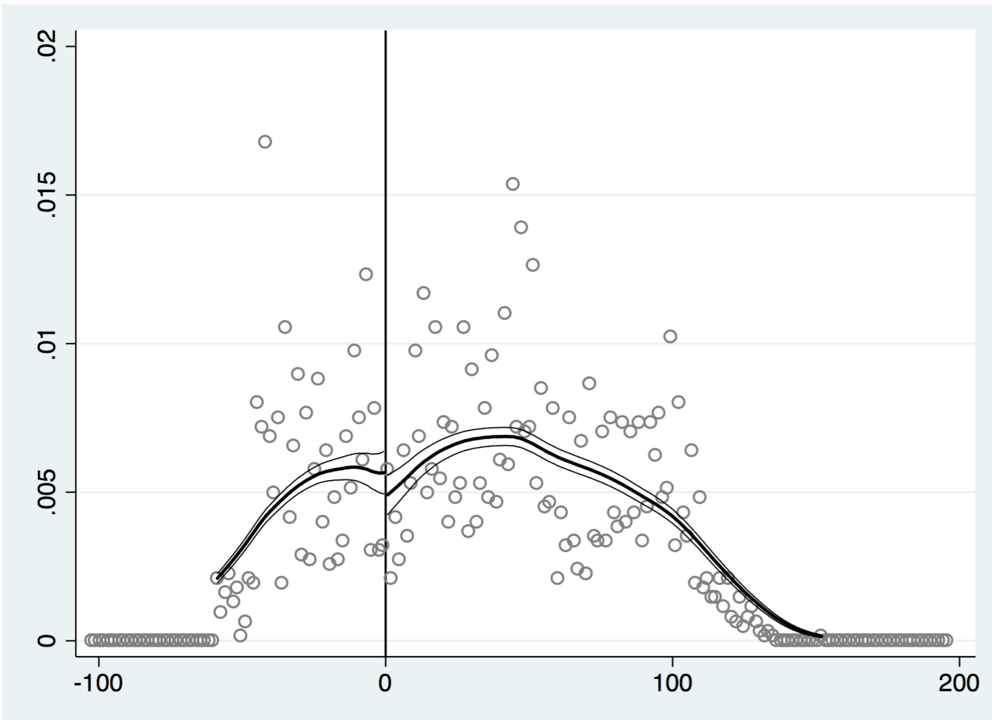
Compared with other Los Angeles Unified teachers on the value-added measure of test score improvement, Kim ranked:

- **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

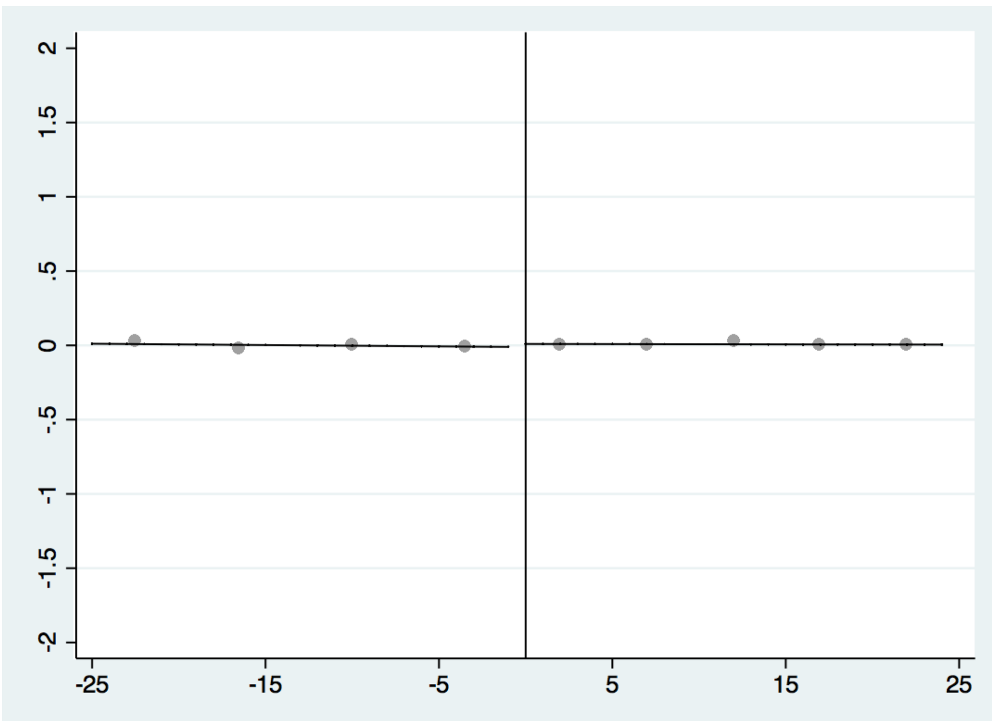
notes: 2009 ranking From LA Times webpage

Figure 2: Density test around cut off



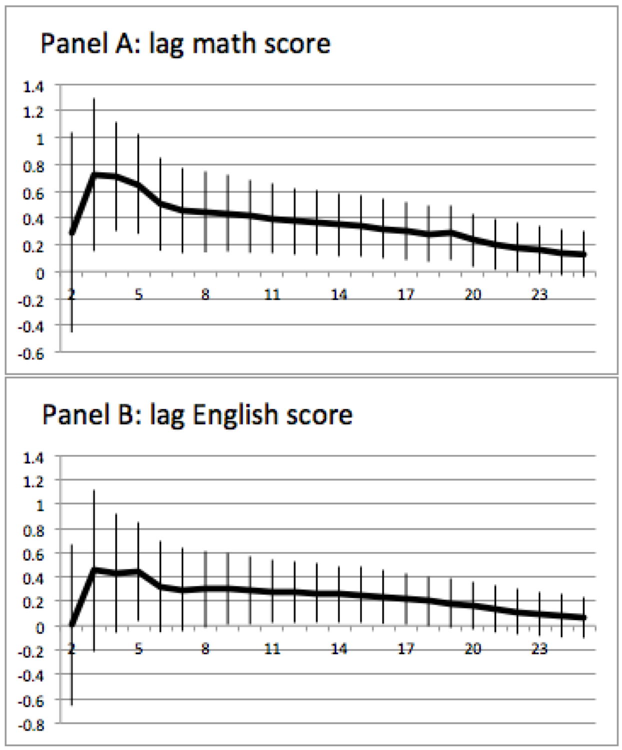
The figure shows the density test proposed by McCrary (2008). We plot the density of observations by the assignment variable (number of students taught with a test score and a lag score). We have re-centered the distribution so that 0 corresponds to the 60 student cutoff point used to determine publication.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

Figure 3: Teacher quality above and below the publication cutoff



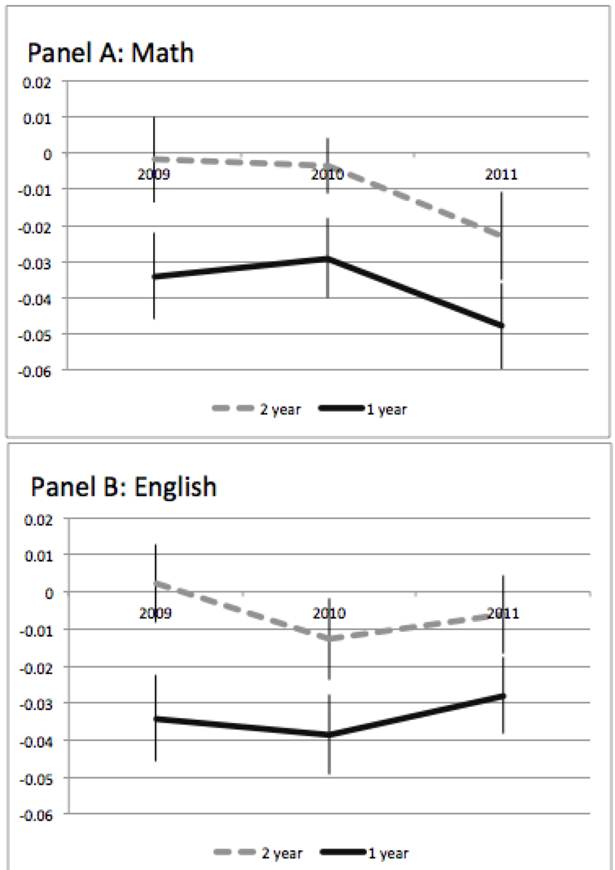
The figure plots the average value added score of teachers above and below the 60 student cutoff. Baseline value added scores are from the year 2010. Value added scores and the number of students per teacher are from the Los Angeles Times.

Figure 4: Coefficient Estimates of Published interacted with Value-Added by bandwidth



The figure shows the coefficient on published interacted with value-added where the dependent variable is lagged math score of students in panel A and lagged English score in panel B. The specification is the one used in Table 3.. Error bars depict a 95 percent confidence interval.

Figure 5: The effect of Value Added Scores after 1 and 2 years



The figure shows the coefficient estimate and a 95 percent confidence interval for β from the following regression: $TestScore_{ij} = \alpha + \beta * year_t * VA_{jt-x} + \gamma lagTestScore_i + X_j'\delta + \epsilon_i j$. Where $TestScore_{ijt}$ is a student is test score with teacher j. $year_t$ is an indicator variable equal to 1 in year t and VA_{jt-2} teacher js value added score from 2003 to $t-x$, where x is 1 for the black line and x is 2 for dashed grey line. Other controls: year fixed effects, grade level fixed effects, cumulative students, cumulative students squared, years in district fixed effects, lagged test score, parents education fixed effects, grade level fixed effects and teacher fixed effects. Sample is all students in year t and $t-1$. Standard errors are clustered at the teacher level.

Table 1: Summary Statistics

Variable	
Students	
Math Score	383.4
	(88.1)
Lag Math Score	379.3
	(68.4)
English Score	347.5
	(59.0)
Lag English Score	350.0
	(58.9)
Parents Education (yrs)	13.5
	(1.32)
Observations	146,001
Teachers	
Years in District	13.7
	(6.8)
Years Teaching	14.1
	(7.4)
White	36.4%
Black	10.4%
Asian	8.6%
Hispanic	36.7%
Female	69.9%
Published in 2010	69.7%
Observations	4,431

Notes: Students are third, fourth and fifth graders in the 2010-2011 school year. Teachers are teachers whom the LA Times calculated value-added and are in the 2010-2011 school year. Standard deviations in parentheses.

Table 2: Teacher Covariate Balance

Variable	Difference	S.E.	p-value	Obs.
Years in District	0.168	0.389	0.67	945
Years Teaching	0.325	0.441	0.46	945
White	0.063	0.028	0.03	1020
Black	-0.012	0.019	0.52	1020
Asian	-0.009	0.018	0.61	1020
Hispanic	0.016	0.031	0.61	1020
Education	0.098	0.068	0.15	945
Female	0.023	0.028	0.41	1020

Notes: Difference is between published teachers with 60 to 79 students and teachers with 40 to 59 students.

Table 3: Student Effects: Regression Discontinuity Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.	**Main Effects (2010-2011)**					
	Lagged Math Scores		Lagged English Scores		Class Size	
Publish	0.044	0.061	0.097	0.082	1.32	1.28
	(0.108)	(0.107)	(0.109)	(0.163)	(0.81)	(0.83)
Publish*Teacher Quality		0.26**		0.18*		-0.26
		(0.10)		(0.10)		(0.83)
Observations	805	805	805	805	929	929
Panel B.	**Placebo (2009-2010)**					
	Lagged Math Scores		Lagged English Scores		Class Size	
Publish	0.149	0.123	0.069	0.052	0.64	0.75
	(0.102)	(0.105)	(0.101)	(0.106)	(0.84)	(0.86)
Publish*Teacher Quality		-0.108		-0.044		1.23
		(0.109)		(0.107)		(0.88)
Observations	868	868	888	888	989	989

Notes: This table shows the coefficient estimates from equation 1. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the school year 2010-2011 with 40 to 79 students previously. All regressions include the teacher's 2003-2010 value added score to control for teacher quality.

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

Table 4: Student Response: Regression Discontinuity Estimates (Distance to Peers)

	(1) Lagged Math test scores	(2) Lagged English test scores
Publish	0.067	0.091
	(0.116)	(0.128)
Publish*Teacher Quality	0.09	-0.051
	(0.143)	(0.155)
Publish*Distance to Peers	0.34***	0.33***
	(0.10)	(0.12)
Observations	789	746

Notes: Table shows the coefficient estimates from equation 1. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the school year 2010-2011 with 40 to 79 students previously. All regressions include the teacher's 2003-2010 value added score to control for teacher quality.

Table 5: Teacher Effects: Teacher Attrition and School Switching

	(1)	(2)	(3)	(4)
Panel A.	**Attrition from Data**			
	In 2011 Data		In 2012 Data	
Publish	-0.052	-0.056	-0.031	-0.026
	(0.062)	(0.062)	(0.070)	(0.071)
Publish*Teacher Quality		-0.329		-0.101
		(0.292)		(0.343)
Observations	989	989	989	989
Panel B.	**Changed Schools**			
	Δ In 2011 relative to 2010		Δ In 2012 relative to 2010	
Publish	-0.035	-0.037	-0.064	-0.057
	(0.054)	(0.053)	(0.072)	(0.073)
Publish*Teacher Quality		-0.225		-0.158
		(0.272)		(0.349)
Observations	989	989	989	989

Notes: Using 2010 as the base year, this table shows attrition from the data and whether teachers switch schools. Columns (1) and (2) shows teacher attrition or teacher switching from 2010 to 2011, and Columns (3) and (4) show teacher attrition or teacher switching from 2010 to 2012. Robust standard errors in parentheses. Sample is teachers in the school year 2010-2011 with 40 to 79 students previously. Teacher quality is based on the LA Times 2003-2010 value added score.

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

Table 6: Effect of Publication on Student Test Scores

	(1)	(2)	(3)	(4)
Panel A	Math		English	
Pub*2011	-0.017	-0.020	0.003	-0.002
	(0.015)	(0.014)	(0.011)	(0.011)
Observations	219,992	257,891	223,600	262,945
Teacher Fixed Effects	yes	yes	yes	yes
Balanced Panel	yes	no	yes	no
Panel B	Math		English	
Share Published	-0.017	-0.019	-0.006	-0.004
	(0.022)	(0.022)	(0.017)	(0.016)
Observations	219,992	257,891	223,600	262,945
Teacher Fixed Effects	yes	yes	yes	yes
Balanced Panel	yes	no	yes	no

Notes: Table shows the coefficient estimates from equation 2. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the 2009, 2010, and 2011 school years. Panel B uses the share of published teachers at the grade level of the explanatory variable of interest.

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

Table 7: Effect of Publication on Student Test Scores

	(1)	(2)	(3)	(4)
Panel A	**Performance Effects 2009-2011**			
	Math		English	
pub*score*2011	-0.025***		-0.034	
	(0.007)		(0.031)	
Least effective		0.009		0.021
		(0.020)		(0.017)
Less effective		-0.002		0.011
		(0.019)		(0.015)
Average		-0.020		-0.009
		(0.019)		(0.015)
More effective		-0.003		0.006
		(0.019)		(0.015)
Most effective		-0.064***		-0.004
		(0.020)		(0.016)
Observations	158,084	158,084	161,328	161,328
Panel B	**Performance Effects 2009-2012**			
	Math		English	
Pub*Score*Year	-0.044***		-0.033***	
	(0.005)		(0.004)	
Least Effective		0.046***		0.052***
		(0.017)		(0.013)
Less Effective		0.021		0.035***
		(0.016)		(0.012)
Average		-0.002		-0.001
		(0.016)		(0.012)
More Effective		-0.027*		-0.001
		(0.016)		(0.013)
Most Effective		-0.069***		-0.036***
		(0.017)		(0.013)
Observations	208,175	208,702	213,033	213,600
Teacher F.E.	yes	yes	yes	yes

Notes: Table shows the coefficient estimates from equation 2. In columns 1 and 3 there is an additional interaction term between a published indicator variable, the published teacher value-added score in that year and year indicator equal to 1 if the teacher was published in that year. In columns 2 and 4 we replace the difference-in-difference variable with indicator variables for quintiles interacted with a year dummy equal to 1 if the teacher was published in that year. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the 2010, and 2011 school years for Panel A and 2010, 2011, and 2012 school years for panel B.

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

Table 8: Mean reversion in 2009-2010

	Math		English		2011 Math	2011 Eng
	(1)	(2)	(3)	(4)	(5)	(6)
pub*score*2010	-0.032***		-0.040***		0.004	-0.004
	(0.006)		(0.006)		(0.006)	(0.006)
Least effective		0.058***		0.040***		
		(0.018)		(0.015)		
Less effective		0.011		0.020		
		(0.018)		(0.015)		
Average		0.001		-0.002		
		(0.019)		(0.015)		
More effective		-0.041**		-0.018		
		(0.018)		(0.015)		
Most effective		-0.027		-0.057***		
		(0.019)		(0.016)		
Observations	163,090	163,090	165,241	165,241	163,090	165,241
Teacher F.E.	yes	yes	yes	yes	yes	yes

Notes: Table shows the coefficient estimates from equation 2 where 2010 is treated as the publication year. In columns 1 and 3 there is an additional interaction term between a published indicator variable, the 2010-2011 published teacher value-added score and year indicator equal to 1 if the year was 2010. In columns 5 and 6 the published score in 2011-2012 is used. In columns 2 and 4 we replace the difference-in-difference variable with indicator variables for quintiles interacted with a year indicator equal to 1 if the year was 2010. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the 2009 and 2010 school years.

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

Table 9: Persistance of effects

	Math score (t+1)	English score (t+1)
	(1)	(2)
pub*score*2011	0.009	-0.005
	(0.007)	(0.007)
Observations	100,705	99,963

Notes: Table shows the coefficient estimates from equation 2 with an additional interaction term between a published indicator variable, the published teacher value-added score in that year and year indicator equal to 1 if the teacher was published in that year. The dependent variable is test score in 2012. Sample is teachers in the 2010, and 2011 school years. Standard errors (in parentheses) are clustered by teacher.

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

Table 10: Achievement Gap

	Math score (1)	English score (2)
Share Published	-0.009	-0.002
	(0.022)	(0.017)
Share Published*lag score	-0.038***	-0.017***
	(0.006)	(0.006)
Observations	219,992	223,600

Notes: Table shows the coefficient estimates from equation 2. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the 2009, 2010, and 2011 school years.

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

Table 11: Top Coding

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. Var=Top Coded				Dep. Var=test score			
	Math		English		Math		English	
	2011	2012	2011	2012	2011	2012	2011	2012
Pub*Score*Year	0.003*	0.002**	0.0004	0.0002	-0.023***	-0.043***	-0.005	-0.032***
	(0.002)	(0.001)	(.0004)	(.0003)	(0.006)	(0.013)	(.0006)	(.004)
Observations	158,084	213,496	161,328	219,199	154,549	207,775	161,028	212,583

Notes: Table shows the coefficient estimates from equation 2 with an additional interaction term between a published indicator variable, the published teacher value-added score in that year and year indicator equal to 1 if the teacher was published in that year. The dependent variable in columns (1)-(4) is an indicator variable for a top coded student. Standard errors (in parentheses) are clustered by teacher. Sample is teachers in the 2010, and 2011 school years for columns (1), (3), (5), and (7) and 2010, 2011, and 2012 for columns (2), (4), (6), and (8). In columns (5)-(8), top coded students are removed from the sample.

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