The Effects of Performance-Based Compensation: Evidence from Oregon's Teacher Incentive Fund

Thomas Dee
Alex Smith
James Wyckoff

ABSTRACT

We examine the impact of systems of differential teacher pay based on student test performance and classroom observations implemented by school districts across the country as part of the Teacher Incentive Fund (TIF), a large competitive grant program operated by the U.S. Department of Education. We use student and teacher-level data from six TIF-participating districts in the state of Oregon to provide early evidence on how these high-profile reforms influenced measures of student performance and teacher retention. We estimate these impacts of TIF by leveraging the discontinuous rule that defines a school's TIF-eligibility in a regression-discontinuity design. Our approach provides an effective proof of concept for TIF in two ways. First, we capture much of the treatment contrast between implementing the full bundle of TIF-reforms and the status quo. Second, we focus on districts with high fidelity of implementation unlike many of the participating districts. We find significant effects of TIF in improving student reading achievement and reducing teacher retention, but we find no effect on math achievement.

Acknowledgement: This work was supported by the Smith Richardson Foundation through grant #R305B090002. The findings and conclusions presented here are mine alone.

1. INTRODUCTION

There are substantial challenges to designing compensation systems to identify and retain workers; these challenges are only compounded for teachers. This is, in part, because effective teaching is multi-dimensional and, in some cases, only realized over students' longer-term development. A further complication is that the identification and reliable measurement of effective teaching practice is still developing and is likely to depend on the specific context.

In most current practice, assessments of public-school teachers tend to be informal, uninformative (e.g. binary ratings of satisfactory or unsatisfactory), and unconnected to both professional-development activities and long-term retention decisions (Weisberg et al, 2009). In fact, over 99% of teachers in districts with these sorts of evaluations were deemed satisfactory (Weisberg et al, 2009). Meanwhile, a number of studies have found substantial variation in teacher quality, suggesting that replacing a 25^{th} percentile teacher with a 75^{th} percentile teacher would increase math test scores by roughly 0.2 standard deviations (Hanushek and Rivkin, 2012). There is also a growing realization that this large variance in teacher quality has economically meaningful implications for students' long-run outcomes (Lazeaer, 2003; Hanushek and Zhang, 2009; Hanushek and Woessman, 2009; Chetty et al., 2011). This has motivated recent policy efforts to design and implement new systems that can reliably assess, support, and encourage teacher quality. The U.S. Department of Education recently sponsored one of the most high profile and large-scale of these initiatives under the aegis of the Teacher Incentive Fund (TIF).

The TIF program was initially authorized in 2006 to provide support to public schools introducing "performance-based compensation systems" (PBCS). In 2010, the U.S. Department of Education increased the scale of this program, using $600 million in stimulus funds to make awards to over 60 multi-district grantees implementing reforms in more than one thousand schools. The revised federal guidance associated with these 2010 TIF awards (referred to as Cohort 3) required that districts implement integrated systems that blend both test-score data and observation-based assessments that guide teachers' professional-development activities as well as decisions about their retention. Additionally, grantees were required to limit the use of TIF funds to high-need schools, defined as those schools with 50 percent or more of their students eligible for free or reduced-price lunches.

We examine the effects of these TIF reforms by focusing on one Cohort 3 TIF awardee, a coalition of six Oregon districts with 131 schools. Using student and teacher-level data, we track outcomes through the TIF planning year and two years of full implementation (three years for teachers). We identify the effects of the TIF reforms through a regression-discontinuity (RD) design that leverages the federally defined eligibility criteria that limited TIF to high-need schools. Across multiple specifications, sample constructions, and bandwidth selections, we consistently find that TIF improved reading achievement significantly in its first year (2011-12) but not in its second year (2012-13). Additionally, we find evidence that TIF also improved teacher retention in 2011-12, suggesting that retention may have been the mechanism for TIF's reading achievement impact.

Our study of this new initiative can be situated in an emerging literature on multi-faceted systems of teacher assessment, feedback, incentives, and support (e.g., Dee and Wyckoff 2013). These more holistic styles of compensation reform pair tools for teachers to improve their

performance with incentives to do so. They stand in contrast to recent narrowly targeted experimental pilots, which provided only cash awards to teachers for test-score improvements and that have generally not been found to increase achievement (Springer et al, 2010; Springer et al, 2012; Fryer, 2013; Glazerman and Seifullah, 2012).

The rest of this paper is organized as follows. In Section 2, we review the economic literature on PBCS. In Section 3, we provide some background detail on TIF. In Section 4, we describe our data. In Section 5 we discuss our research design. In Section 6, we present our findings on student achievement effects, and in Section 7, teacher retention effects. Finally, in Section 8, we provide a concluding discussion.

2. EXISTING EVIDENCE ON PERFORMANCE-BASED COMPENSATION SYSTEMS

There is a large, decades-old literature that explores variants of performance-based compensation systems (PBCS) for teachers. However, until recently, the methodological rigor of the available empirical evidence has generally been weak.[1]

Most PBCS are designed to improve student achievement (1) by providing financial incentives for teachers to exert more effort or develop stronger teaching skills, and (2) by increasing incentives for effective teachers to enter and remain in schools. Several recent city-level studies provide evidence solely on the effort margin by examining the extent to which the productivity of existing teachers increases in "cash for test score" incentive schemes. First, the Project on Incentives in Teaching (POINT) was a 3-year study that provided randomly assigned middle-school mathematics teachers in Nashville bonuses of as much as $15,000 if their students

[1] For an overview of this literature, see Springer (2009) or Johnson and Papay (2009).

met ambitious performance thresholds (Springer et al. 2010). The availability of these incentives led to no detectable effects on measured student performance or on measures of teacher effort and classroom practice. A second random-assignment study provided New York City teachers with rewards up to $3,000 for meeting performance targets (Fryer, 2013). In this study, treatment schools had flexibility in designing their incentive schemes, and most chose group-based incentives. The impact estimates from this study suggest that the presence of incentives did not raise school performance and may have even lowered it. A third study was conducted in 34 Chicago schools that were randomly assigned to when (but not if) they implemented the Teacher Advancement Program (TAP). Under this program, which was somewhat broader than a simple cash for test score scheme, teachers were eligible to receive payouts of as much as $6,400 for their contribution to the achievement-based value-added of their students (at the school and school-grade level) and their performance on a classroom observation rubric. Under TAP, teachers could also earn extra pay for undertaking the increased responsibilities associated with promotion to a mentoring or master status. The evidence from this study suggests that random assignment to TAP did not raise student achievement (Glazerman and Seifullah, 2012). However, the program implementation did not occur entirely as intended. Teacher payouts were smaller than the originally stated targets and there were no rewards based on value-added because the requisite linked data systems were inadequate.

In contrast to the null findings in studies of cash for test score incentive schemes, Dee and Wyckoff (2013) find substantial positive effects of incentives on teacher performance and retention in the context of IMPACT, the teacher-evaluation system introduced in the District of Columbia Public Schools. Unlike the incentive schemes previously studied, IMPACT based its high-powered incentives (i.e. dismissal threats and financial rewards) on multiple measures of

teacher performance, including both test performance and structured observational measures. Dee and Wyckoff use a RD design that leverages discontinuities in rewards or threats at performance score thresholds to identify the effects of these incentives. They find that dismissal threats increased the voluntary attrition of low-performing teachers more than 50 percent and improved the performance of the low-performing teachers who remained by 0.27 of a teacher-level standard deviation. They also find evidence that financial incentives further improved the performance of high-performing teachers (effect size = 0.24).

Aside from the motivational impacts of incentives, PBCS advocates argue that these systems have the potential to attract and retain a different type of teacher, one who expects to realize the compensation benefits of meeting performance targets (Lazear, 2003). These advocates fault rigid compensation systems that reward only experience and education for the sharp decline in the aptitude of those selecting into the teaching profession over the last 40 years (Corcoran et al. 2004, Hoxby and Leigh 2004). The potential of PBCS to impact such patterns of self-selection into and out of the teaching profession may be a particularly important mechanism for improving overall teacher quality. However, we are aware of relatively little empirical research that explores this important aspect of PBCS directly. One example is a panel-based evaluation of a targeted teacher bonus in North Carolina (Clotfelter et al. 2008), which finds that a $1,800 bonus reduced teacher turnover by 17 percent. The TAP evaluation described above also found limited evidence that incentives improved teacher retention (Glazerman and Seifullah 2012). Another example is the Dee and Wyckoff's study of IMPACT, which found dismissal threats increased the likelihood that teachers left the district. However, the field-experimental studies in New York and Nashville provide no evidence that compensation incentives influenced teacher retention (Springer et al. 2010, Fryer et al. 2011).

Our contributions speak to the evaluation of the resource-intensive TIF program and the more general effectiveness of PBCS that employ incentives based on multiple measures (i.e. classroom observations and school value-added) in combination with high quality evaluative feedback and professional development. First, we analyze a program which differs markedly from most other PBCSs that have been studied because of its more holistic approach. Second, we provide direct, timely evidence on the potential effectiveness of a future expansion of the TIF program to less needy schools by estimating the effect of TIF *at* the high-need threshold. Third, our credible evaluation of the student performance effects of the TIF program is important simply because the program is an unusually large-scale and exceptionally expensive federal effort (i.e., $1.2 billion over five years to 62 multiple-partner programs in 27 states).

3. TEACHER INCENTIVE FUND (TIF)

In this section, we provide institutional details on the Teacher Incentive Fund (TIF). First, we describe the evolution of the federal grant program. Then, we discuss the specifics of the implementation of TIF in Oregon, focusing in particular on the contrast between TIF and non-TIF schools.

3.1 Federal Competitive Grant Program

The U.S. Department of Education's (USDOE) TIF was initially authorized with $99 million in federal appropriations in 2006.[2] Its goal was to support education agencies that develop and implement performance-based compensation systems (PBCS) in primary and

[2] "Teacher Incentive Fund: Funding," http://www2.ed.gov/programs/teacherincentive/funding.html (October 8, 2014).

secondary education settings. In 2007, 33 grants were given to state or local education agencies in 18 states (cohorts 1 and 2).[3] In 2010, dramatically increased funding from the American Recovery and Reinvestment Act of 2009 ($200 million) and the Consolidated Appropriations Act of 2010 ($400 million) permitted a larger third cohort of 63 grants to agencies in 27 states.[4] For Cohort 3, USDOE imposed a number of new, explicit requirements to grant applicants. First, applicants were required to develop and implement a PBCS that rewards both teachers and principals who demonstrate their effectiveness in improving student achievement. PBCSs had to give significant weight to student growth based on test score data (with a preference given to value-added models) and also include observation-based assessments. They also had to provide substantial incentive amounts to teachers and principals. Second, applicants were required to present an integrated strategy that incorporated test score data and observation-based assessments in professional development as well as retention and tenure decisions. Third, applicants were directed to demonstrate the fiscal sustainability of their PBCS and accept responsibility for funding an increasing share of the PBCS using non-TIF sources. Fourth, and particularly important for our methodological design, TIF funds could only be used to fund the incentive awards and professional development of teachers and principals in high-need schools. These eligible schools were defined as those where free and reduced price lunch eligible (FRL) students made up at least 50% of their enrollment in 2009-10.[5]

3.2 Oregon TIF

[3] "Teacher Incentive Fund: Awards," http://www2.ed.gov/programs/teacherincentive/2007-awards.html (June 11, 2012).
[4] "Teacher Incentive Fund: Funding," http://www2.ed.gov/programs/teacherincentive/funding.html (October 8, 2014).
[5] Under the Richard B. Russell National School Lunch Act, students from families below 185% of the federal poverty level are eligible for free or reduced price lunch.

In 2010, the Chalkboard Project, an Oregon-based non-profit organization in partnership with six Oregon school districts, was awarded $24.4 million over five years by the Teacher Incentive Fund.[6] After a planning year (2010-11), the TIF implementation began in the 2011-12 school year in the selected TIF schools in each district.[7] Not all TIF-eligible schools (those with at least 50% FRL) were selected as TIF schools. Among their TIF-eligible schools, most districts gave preference to those schools that were underperforming in at least one subject compared to similar schools in the state. Of the 147 schools in the six participating districts (Bend-La Pine, Crook County, Greater Albany, Lebanon, Redmond, and Salem-Keizer), 41 schools received the TIF performance-based awards (TIF schools).[8]

Prior to TIF, these six districts had received CLASS (Creative Leadership Achieves Student Success) grants of up to $30,000 (as well as up to $30,000 of in-kind technical support) from the Chalkboard Project to offset the costs of designing and planning a framework to integrate performance evaluation, relevant professional development, and expanded career opportunities for teachers and principals.[9] Due in part to CLASS, much of each district's policies around career paths, performance evaluation, and professional development were the same for TIF and non-TIF schools. However, important differences remained, which together constitute the TIF "treatment" in Oregon. First, teachers in TIF schools received incentive awards based on their performance, while teachers in non-TIF schools did not. The details of the award models differed by district, but all districts provided awards of up to approximately $6,000

[6] Initial award of $13.2 million was later increased to $24.4 million.
[7] The first performance awards were made during the 2012-13 school year using results from the teacher evaluations for 2011-12.
[8] In two districts, Bend-Lapine and Greater Albany, 22 of the 25 schools selected to participate in TIF also participated in the national TIF evaluation. This meant that these schools were randomly assigned to either be TIF schools or control schools, where the control condition involved a stipend given to all teachers rather than performance-based awards. In all, 11 schools were in this control group.
[9] "Request for Proposals, CLASS Project Pilot Program", The Chalkboard Project (February 2012).

for teachers ($12,000 for principals) based on school-level value-added measures and teacher-level (or principal-level) evaluations.[10] Second, only teachers and principals in TIF schools could receive TIF-funded professional development.[11] This restriction was important because CLASS did not include implementation funding to ensure the adoption of new policies, such as increased professional development, district-wide. This made professional development for teachers in TIF schools easier for districts to provide, but it also resulted in higher quality teacher evaluations (i.e. more observations and more feedback) by administrators in TIF schools, according to district interviews conducted by the Chalkboard Project.[12] These higher quality evaluations were a result of TIF school administrators receiving substantially more training than their non-TIF counterparts. For example, they were required to complete Teachscape, an evaluator training program which includes scoring videos of classrooms and two multiple hour tests; they also attended a 3-day summer institute which included training on how to give effective feedback.[13]

4. DATA

In this section, we discuss the data used in our empirical analysis (subsection 4.1), the construction of the assignment variable in our regression discontinuity design (subsection 4.2), and the selection of our analytical sample (subsection 4.3).

4.1 Student Achievement Data

[10] "Grantee Profiles: Chalkboard Project," http://www.cecr.ed.gov/TIFgrantees/granteeProfiles/cohort3.cfm?id=9 (March 30, 2012).
[11] "Teacher Incentive Fund: Frequently Asked Questions for the 2010 Competition and Grant Awards," http://www2.ed.gov/programs/teacherincentive/2010-faqs.doc (June 28, 2010).
[12] B. Pratt (personal communication , June 4, 2014).
[13] B. Pratt (personal communication , June 4, 2014).

We obtained student-level achievement data for all public schools in Oregon for school years 2007-08 through 2012-13 from the Oregon Department of Education through a restricted-use data agreement. Observations in this data include all of the students who took the math and reading Oregon Assessment of Knowledge and Skills (OAKS). The OAKS is taken by all students in Oregon public schools in grades 3-8 and then once in high school. It measures student performance against the Oregon Content Standards and assigns raw scores between 150 and 300. The data contain student and school identifiers, student grade level, raw test scores, and student demographic information (e.g. free and reduced-price lunch eligible, ethnicity, and sex). We normalize the raw test scores by subtracting the state-wide mean score for each grade, subject, and year combination and then dividing by the state-wide standard deviation for that test-subject-year.[14] The resulting standard deviation scores allow for easy comparisons across grades, years, and subjects.

4.2 Teacher Data

We also obtain data for all teachers in Oregon public schools for school years 2006-2007 to 2013-2014 from the Oregon Department of Education. This data contains teacher and school identifiers, full-time equivalency (FTE), years of experience, grade levels taught, and demographic information (e.g. birth date, sex, ethnicity, educational attainment). The raw teacher data is observed at the teacher-school-position-year level. We collapse it to the teacher-year level, summing FTE across positions in a given teacher-year, and we restrict our analysis sample

[14] For students with multiple test records on the same test in the same day likely as a result of a data entry error, we use the maximum score. For students that took the same test multiple times on different days, we use the score from their most recent test.

to only full-time teachers.[15,16] Unfortunately, we are not able to use this data to match teachers to their students, only to their schools. This makes it impossible to calculate individual teacher value-added measures.

The primary focus of our analysis of TIF's impact on teachers is retention. We operationalize this by constructing an indicator of teacher exit. For full-time teachers the year prior to implementation (2009-10 or 2010-11, depending on our definition of implementation), this is defined as equal to one for those that are no longer full-time teachers in the same school in the given observation year (2010-11, 2011-12, 2012-13, or 2013-14).

4.3 Assignment Variable Construction

Our RD design identifies the effect of TIF by comparing students in schools that were just eligible for TIF with those that were just ineligible. Consequently, the measure by which this eligibility is determined, the assignment variable, is critical to our empirical approach. It is intended to reflect the eligibility of a school at the time when the school's district made the decision to select or not select it as a TIF school. While we were unable to obtain the precise data used by the districts for their TIF selection, we recreated this data as accurately as possible given Oregon Department of Education and Chalkboard Project documentation and in consultation with representatives from those organizations.

According to USDOE's eligibility criteria, a school could receive TIF funding if at least 50% of students were eligible for free and reduced price lunch (FRL) .[17] Middle and high

[15] Approximately 4% of teacher-years are associated with multiple schools (i.e. teachers hold positions in more than one school in a given year). In these cases, I keep only school observation for the school where the teacher spent the majority of their time (highest FTE) in the given year (and that FTE is greater than 50%).
[16] Full-time teachers are defined as those with a FTE of at least 0.9.
[17] This eligibility restriction applies only to funding for incentive awards and professional development. More general funding, such as for data system upgrades could benefit schools that were not "TIF-eligible".

schools could also be deemed TIF-eligible if their "feeder schools" exceeded the 50% FRL threshold.[18] We construct an assignment variable z_s for a given school s that represents how close the school was to satisfying these eligibility criteria at the time that TIF schools were selected by the participating districts. We define this assignment variable as follows:

$$z_s = \max(FRL_s, FEEDFRL_s) - 50,$$

where school *s* was TIF-eligible if and only if $z_s \geq 0$. FRL_s is the percent of the enrollment of school s that was eligible for FRL at the time of selection. For middle and high schools, $FEEDFRL_s$ is the percent of the enrollment that was eligible for FRL in schools that fed into s. For elementary schools, $FEEDFRL_s$ is set to zero so that it has no effect on z_s.

To the extent possible, we use the FRL data that is recorded nearest to the time that final TIF selections were made in the Summer of 2010. For the Bend-Lapine School District, we were able to obtain FRL data from the district recorded in May 2010. For the remaining TIF-participating districts, we were only able to obtain FRL data from the Oregon Department of Education's state-wide database, recorded in Fall 2009.[19,20]

Following Oregon Department of Education guidelines, we define $FEEDFRL_s$ as the average percent FRL of students entering school s from all of its feeder schools.[21] This is constructed by taking the sum of FRL students in all of a school's feeders and dividing it by the

[18] A school's "feeder schools" are those schools that send their graduating students to the school. If 5th grade students from elementary schools A, B, and C go to middle school D as 6th graders, then A, B, and C are feeder schools for D.
[19] If Fall 2009 state data is used for Bend-Lapine, not all TIF schools are
[20] For the one school in TIF-participating disctricts that was opened in the 2010-11 school year, Rosland Elementary, we use FRL data for the 2010-11 school year (recorded in Fall 2009) from the Oregon Department of Education's state-wide database.
[21] Janet Bubl. 2014. "Establishing Poverty Levels." Oregon Department of Education Memo.

sum of enrollment in all those feeders, where each sum is weighted by the fraction of students coming from each feeder school. This can be written as follows:

$$FEEDFRL_s = 100 \cdot \frac{\sum_{f \in F_s} p_{sf} \cdot NFRL_f}{\sum_{f \in F_s} p_{sf} \cdot NENROLL_f}$$

where F_s is the set of feeder schools for school s. $NFRL_f$ is the count of FRL students in feeder school f and $NENROLL_f$ is the count of all students in feeder school f.[22] p_{sf} is the fraction of students at school s that came from feeder school f.[23] Figure 1 shows the school-level distribution of z_s for the analytical sample. Few schools (and students) have z_s below -25 (25% FRL) or above 40 (90% FRL).

4.4 Analytical Sample

We make a number of restrictions to the universe of Oregon public schools to obtain our analytical sample. First, we restrict to schools that were in existence in 2010-11, the year before the implementation of TIF, in one of the six TIF-participating districts (145 schools). This ensures that our analysis only compares schools in districts that chose to apply for the TIF grant with Chalkboard. Second, we exclude non-traditional schools (8 charter schools and 4 K-2 schools), leaving a sample of 133 schools (80 elementary schools, 22 middle schools, 17 high schools, and 14 schools with other grade configurations). Under our definition of the assignment variable, the sample can be categorized into three groups: 41 schools were TIF-ineligible ($z_s < 0$) and *did not receive* the TIF treatment, 51 schools were TIF-eligible ($z_s \geq 0$) and *did not*

[22] When calculating $FEEDFRL_s$ for a high school, if a feeder school f is a middle school then its own feeder FRL and enrollment is used, $NFRL_f = \sum_{g \in F_f} p_{fg} * NFRL_g$ AND $NENROLL_f = \sum_{g \in F_f} p_{fg} * NENROLL_g$.

[23] We calculated p_{sf} using unduplicated counts of students by school attended 10/1/2011 and school attended 10/1/2010, obtained from Oregon Department of Education. We restricted this data to transitions representing students graduating from one school and entering another school within the same district (e.g. 5th grade to 6th grade transitions). p_{sf} was then defined as the fraction of students entering school s that had been enrolled in feeder school f in the prior year.

receive the TIF treatment, and 41 schools were TIF-eligible and *received* the TIF treatment.[24] Among the TIF-eligible schools ($z_s \geq 0$), the decision to assign schools to the TIF treatment was made by districts, often based on whether they had underperformed in the previous year (relative to a comparison group of schools in the state).[25] Table 1 shows summary statistics for schools in the analytical sample in 2009-10. Tests scores are observed for 44,624 students, their average in math and reading was slightly below the state average. Students in TIF-ineligible schools were more likely to be white, less likely to be economically disadvantaged, and averaged much high test scores (0.3-0.4 SD) than students in TIF-eligible schools. Students in TIF schools were similar in demographic composition to students in TIF-eligible non-TIF schools, but averaged lower test scores (0.09 SD).

5. EMPIRICAL STRATEGY

We evaluate the effects of the stimulus-funded TIF programs in a fuzzy regression discontinuity (fuzzy RD) design that leverages the TIF-eligibility requirement mandating that these reforms be targeted only to high-need schools. The basic intuition behind the strong causal warrant of this RD design is straightforward. Within the districts receiving TIF awards, schools that just barely meet the high-need standard are much more likely to implement TIF. In contrast, schools that just fail to meet the high-need standard are ineligible for TIF. That is, whether a school is just above or below this threshold implies a strong contrast in the PBCS regime that is experienced. This localized variation in whether a school just meets or fails to meet this condition can be viewed as effectively random so long as it is not manipulated, an issue that we

[24] The list of TIF schools was obtained from the Chalkboard Project. "TIF treatment" schools exclude TIF schools randomly assigned to the control group as part of the national evaluation of TIF.
[25] B. Pratt (personal communication , June 4, 2014).

will examine below. Any differences in outcomes among students (or teachers) in schools just above and below the high-need threshold can therefore be credibly attributed to the effects of TIF eligibility and use. Consequently, the RD design alleviates concerns that treated and non-treated groups differ in unobservable ways.

5.1 Regression Discontinuity Specification

We examine this treatment contrast through first-stage specifications that take the following general form:

$$TIF_{is} = \alpha I(z_s \geq 0) + f(z_s) + \theta X_{is} + \epsilon_{is}, \qquad (1)$$

where α identifies the discrete change in TIF-funded reforms experienced by students (or teachers) in schools meeting the eligibility threshold (i.e. $I(z_s \geq 0)$). This identification conditions on how the probability of undertaking TIF reforms relates to the assignment variable, $f(z_s)$ and to control variables reflecting the observed traits of a given student (or teacher) *i*, X_{is}.

A similarly structured reduced-form (or intent-to-treat) specification is applied to the outcome measures, Y_{is}. In our student analysis, Y_{is} is the math and reading achievement (i.e. normalized OAKS score) for student *i*. In our teacher analysis, Y_{is} is the indicator for school exit of teacher *i*. We define our primary measure of school exit as one if teacher *i* is no longer a full-time teacher at the same school as in the pre-implementation year (2009-10 or 2010-11 depending on how it is defined).[26] We define our secondary exit measure, year-over-year school exit, as one if teacher *i* in year *t* is not a full-time teacher in the same school that they were in year *t-1*. The specification is as follows,

[26] For the sample of full-time teachers in TIF-participating districts (excluding charter schools).

$$Y_{is} = \gamma I(z_s \geq 0) + \kappa(z_s) + \beta X_{is} + \nu_{is}. \quad (2)$$

That is, estimates of γ identify the discrete shift in outcomes associated with being in a TIF-eligible school.

We examine the causal warrant of this method by performing a series of tests commonly used when implementing RD designs as well as tests made possible by the panel-nature of our data. First, we restrict the analytical sample to include only students (or teachers) in schools with values of z_s close to the threshold (e.g. plus or minus 20 percentage points). Second, we provide falsification tests that look at the estimated placebo effect of TIF prior to its implementation. Third, we check for covariate balance around the threshold by estimating Equation 2 with student (or teacher) covariates as the outcome variable. Fourth, following McCrary (2006), we test for manipulation of the assignment variable by investigating whether there is a discontinuity in the density of schools at the eligibility cutoff. This test is done at the school-level rather than the student-level, as the school is the unit that could conceivably manipulate their FRL in order to be eligible for TIF. Manipulation is unlikely, however, since in most cases it would require schools to anticipate their district's TIF award more than half a year before the application was sent.

In an additional analysis, we alter our assignment of students to schools to address concerns that changes in student mobility (e.g. more talented students switching to TIF schools) may be driving estimated TIF effects. In this analysis, we restrict our analytical sample to students in grade 3-6 in 2009-10 (or alternatively 2010-11) and assign these students to a predicted school in the TIF treatment years based on the most common school in each year for students in a given school-grade in 2009-10 (or 2010-11). This approach also allows us to

control for a student's baseline test score (in 2009-10 or 2010-11). We therefore remove individual school choice as a potential factor, while also controlling for a clean baseline measure of achievement at the individual student-level.

5.2 Difference-in-Regression Discontinuity Specification

We expand on our RD approach to measuring student achievement impacts of TIF using a Difference-in-Regression Discontinuity (DRD) design, which compares the differences between just eligible and just ineligible students across multiple years simultaneously.[27] The primary purpose of this approach is to explicitly test whether discontinuities in the TIF years (2010-11 to 2012-13) statistically differ from any pre-existing discontinuity at the eligibility threshold in the pre-TIF years, which are grouped together (2007-08 to 2009-10). If the discontinuities in TIF years do not differ statistically from pre-TIF years, it would suggest that any effects estimated in the TIF years using Equation 2 are the product of pre-existing differences around the eligibility threshold. The reduced-form specification is as follows,

$$Y_{ist} = \gamma_0 I(z_s \geq 0) + \kappa_0(z_s) + \sum_{\tau=2011}^{2013} I(t = \tau) \cdot \left(\gamma_\tau I(z_s \geq 0) + \kappa_\tau(z_s)\right) + \beta X_{ist} + \nu_{ist}, \qquad (3)$$

where t is the school year (2011 refers to 2010-11), $I(t = \tau)$ is an indicator equal to 1 for TIF school year $\tau \in \{2011, 2012, 2013\}$, and the other variable definitions follow from Equation 2. γ_0 provides the change in achievement at the eligibility threshold in the Pre-TIF years, while γ_τ provides the differential change in achievement in year $\tau \in \{2011, 2012, 2013\}$ compared to the pre-TIF period.

5.3 Interpretation of Regression Discontinuity Estimates

[27] A similar design is used by Fitzpatrick (2010), though she compares birthdate discontinuities across treated and untreated states rather than across treated and untreated years.

If schools were randomly selected into the TIF treatment from among the group of TIF-eligible schools, then the ratio of the reduced-form estimate from Equation 2 and the first stage estimate from Equation 1 $\left(\frac{\gamma}{\alpha}\right)$ gives the local average treatment effect (LATE) of TIF, that is, the average treatment effect of PBCS for students (or teachers) in schools near the eligibility cutoff. However, this was not the case in Oregon, where TIF schools were selected by the district from TIF-eligible schools, often because they had underperformed in the previous year. This means that $\frac{\gamma}{\alpha}$ can be better interpreted as the local average treatment on treated (LATOT) effect. If district administrators selected for TIF treatment the TIF-eligible schools that were most likely to benefit from the treatment, then the LATOT would exceed the LATE. In this case, estimates of $\frac{\gamma}{\alpha}$ can be considered an upper bound of the LATE, and expanding the TIF program to additional eligible schools that were not initially selected would be expected to yield more modest effects.

6. STUDENT ACHIEVEMENT RESULTS

In this section, we first present regression discontinuity (RD) estimates of the impact of TIF on math and reading achievement. Second, we discuss a number of robustness and falsification tests of our main results. Third, we present an alternative RD approach which alleviates concerns that differential student sorting determines our main results. Fourth, we present difference-in-regression discontinuity results that explicitly test differences in the achievement discontinuities in different years. Finally, we summarize our findings.

6.1 Main Results for Regression Discontinuity Design

Table 2A and 2B show the OLS regression discontinuity estimates of α and γ from Equation 1 (first stage) and Equation 2 (reduced-form) for math and reading standardized scores in each TIF-treatment year (2011-12 and 2012-13). The tables include results for specifications with different sets of individual covariates, all with standard errors clustered at the school-level. The estimates for α (first row) show a statistically and economically significant first stage, which does not vary across specifications or years. However, in some specifications, the first stage F-statistic falls slightly below 10, the commonly applied rule of thumb for avoiding a weak instrument problem, which could inflate effect estimates. The magnitude of the estimates suggests that TIF eligibility increases the probability of receiving the TIF treatment by roughly 40 percentage points. Figure 2 depicts the first stage graphically. For 4 percentage point bins of the assignment variable, z_s, the figure shows the proportion of test-takers in TIF-treatment schools, and superimposes the regression lines from Equation 1.[28]

The second row of Table 2A shows the reduced form estimates (γ in Equation 2) for math standard deviation scores in the same years and using the same specifications. Figure 3A depicts these results graphically (equivalent to column 1). These results show no significant effects of TIF eligibility on math standard deviation scores, with the exception of a marginally significant (at the 10% significance level) and unexpectedly negative impact in 2012-13 (column 10) that is not robust across specifications. Table 2B shows positive and statistically significant reduced-form effects of TIF on reading standard deviation scores in 2011-12, but not in 2012-13. This can also be seen graphically in Figure 3B (equivalent to column 1). The estimates for 2011-12 can be interpreted as a 0.10-0.16 standard deviation increase in reading achievement from TIF eligibility. The resulting estimate of $\frac{\gamma}{\alpha}$ is 0.26-0.41 standard deviations that is statistically

[28] Figure 2 uses the sample of reading test-takers, but the graph is identical for math test-takers.

different from zero, but imprecisely estimated. As discussed in Section 5, this magnitude likely represents the upper bound of the local average treatment effect (LATE) rather than the LATE itself.

6.2 Tests of Internal Validity

Table 3A and 3B present robustness checks for the results in Table 2A and 2B by showing the same specifications estimated on a smaller window around the eligibility cutoff (i.e. $z_s \in [-30, 30]$ and $z_s \in [-20, 20]$ or equivalently $\%FRL \in [20,80]$ and $\%FRL \in [30,70]$).[29] The reduced-form results remain qualitatively similar for smaller windows, with two minor exceptions. First, the 2011-12 reduced-form results for reading increase slightly in magnitude for smaller bandwidths. Second, the 2012-13 reduced-form results for reading become statistically significant for two of three specifications when using the $z_s \in [-30, 30]$ bandwidth (equivalent to $\%FRL \in [20,80]$). Figures A1 and A2 depict the reduced-form results for $\%FRL \in [30,70]$ graphically.

Table 4A and 4B provide a falsification test of whether our empirical strategy finds a TIF placebo effect prior to its implementation, in which case the validity of our approach would be doubtful. These tables show estimates of α and γ from Equations 1 and 2, as in Table 2A and 2B, but for the years prior to TIF implementation. The results show small and statistically insignificant reduced-form effects for either math or reading in the years prior to TIF (2007-2008 through 2009-2010) or in the TIF planning year (2010-2011). This evidence lends substantial support to interpreting the estimated achievement effects in Table 2A and 2B as products of TIF

[29] The relatively few number of schools prevents estimating tighter bandwidths.

and not as a consequence of observed or unobserved differences between students on either side of the eligibility threshold.

Table 5 provides tests of whether test-taker covariates differ discontinuously on either side of the eligibility threshold. The existence of such a difference would present the possibility that achievement differences at the eligibility threshold are due to observable differences between schools rather than TIF eligibility. The table includes estimates of γ in Equation 2 for various bandwidths, where the dependent variable is a student characteristic (X_{is} is excluded). The results show no significant discontinuities in the likelihood a test-taker is economically disadvantaged, but does show significant discontinuities in the likelihood a test-taker is white (11-12 percentage points) and special education students (1-2 percentage points, marginally significant in 2012-2013 only).

We assess the importance of this possible covariate imbalance in two ways. First, we note that no discontinuities in covariates remain significant for the smallest bandwidth, suggesting that these significant estimates may be driven by the influence of schools far from the threshold on the slopes of $\kappa(\cdot)$. The significance of the results in Table 3 were not similarly affected by bandwidth restrictions. Second, we construct indexes of the reading and math test scores predicted from a student's covariates and then check for a threshold discontinuity in this index. This is a more appropriate test of imbalance at the threshold, as covariate imbalance is only problematic to the extent that those covariates affect achievement. We find no significant discontinuities in predicted score indexes at the eligibility threshold for any bandwidth size.

Figure 4 shows the results of the McCrary density test at the school-level, which checks for a discontinuity in the distribution of the assignment variable at the eligibility threshold

(McCrary, 2006). Such a discontinuity could be evidence of schools manipulating the assignment variable in order to make themselves eligible for TIF. We find no evidence of a discontinuity, which is not surprising since the assignment variable was measured more than half a year prior to the TIF application for most schools. We address the possibility that students are "manipulating" their assignment variable with their choice of schools in the following subsection.

6.3 Cohort-tracking Regression Discontinuity Results

Table 6A and 6B address the possibility that the estimated TIF effects in Table 2A and 2B are due to changes in student mobility to TIF and non-TIF schools. For example, if relatively high-achieving students are more likely to move to TIF schools, then increases in achievement in those schools could be due to the composition of their students rather than improvements to teacher/principal quality from TIF. These tables provide estimates of α and γ in Equations 1 and 2 on two cohort subsamples (grade 3-6 students in 2009-10 and in 2010-11), where students are assigned to the school in 2011-12 and 2012-13 attended by the majority of students that were in their school and grade in the baseline year.[30] In this way, we ensure that differential sorting by achievement is not driving our estimates. With this construction, we are also able to control for individual-level baseline test scores (in 2009-2010 or 2010-2011), rather than school by grade test score means.

Table 6A shows no statistically significant effects of TIF on math achievement in 2011-12 or 2012-13. In the main math achievement results in Table 2A, only the estimate in 2012-13 that controlled for school by grade test score means (in 2009-10) was marginally significant

[30] Estimates for 2012-13 are only shown for the 2010-11 cohort because the 2009-10 cohort's youngest members are 6th graders in 2012-13 leading to the loss of elementary schools from the sample (half of all schools) and therefore the loss of the school-level variation necessary for the RD design to work.

(column 6). In Table 6A, the individual baseline test score specification (column 6) produces a 2012-13 estimate that is not significant. Table 6B shows large and significant reduced-form estimates of improved reading achievement in 2011-12. In fact, our preferred individual baseline test score specification (column 3) finds results in line with Table 2B. These results imply estimates of $\frac{\gamma}{\alpha}$ of 0.21 to 0.28 standard deviations. There are also large and significant reduced-form estimates for reading in 2012-13, but not for our preferred specification (column 6). Overall, the results in Table 6A and 6B are qualitatively similar to our main results in Table 2A and 2B, suggesting that those main results are not driven by student mobility or insufficiently precise baseline achievement controls.

6.4 Difference-in-Regression Discontinuity (DRD) Results

Table 7A and 7B present DRD estimates of $\gamma_0$and γ_τ from Equation 3 for two school year samples. The first sample (columns 1 and 2) includes school years 2008-09 to 2012-13 and groups the pre-TIF school years together to test whether the achievement discontinuities at the eligibility threshold in the planning and TIF treatment years differ substantially from the prior years. The tables show insignificant estimates of pre-TIF achievement discontinuities (γ_0) in both math and reading. In Table 7A, there are no significant differences between math achievement discontinuities in TIF years and pre-TIF years. In Table 7B, there is a positive difference between the 2011-12 reading achievement discontinuity and the pre-TIF discontinuity that is significant at the 10 percent level when controlling for student covariates in column 2 (column 1, without controlling covariates, finds an estimate significant at the 16 percent level). The magnitude of these differential discontinuity estimates are in line with those in Table 2A, columns 1-3.

The second sample presented in Tables 7A and 7B, columns 3-4, contains only the TIF treatment years (2011-12 and 2012-13). These results provide a test of whether the difference in the achievement discontinuities between 2011-2012 and 2012-2013 are statistically significant. For both math and reading achievement, estimates of the "ELIGIBILITY X (SY2012-13)" coefficient (γ_{2013}) are not significant, suggesting that while the differences in magnitude between estimated effects in 2011-12 and 2012-13 may be meaningful from a policy stand-point, they are not statistically meaningful.

6.5 Summary of Results

Across multiple specifications, sample constructions, and bandwidth selections, we consistently find that TIF improved reading achievement significantly in its first year (2011-12) but not in its second year (2012-13). We also find that the estimated effect in 2011-12 differs statistically from the Pre-TIF placebo effect. The preferred reduced-form estimates of these reading improvements (Table 2B, columns 3 and 6) are 0.1 SD in 2011-12 and 0.05 SD in 2012-13, implying LATOT effects of 0.26 SD in 2011-12 and 0.11 SD in 2012-13. While the differences in these magnitudes are clearly important, they are not statistically different due to imprecision in the estimates (Table 7B, columns 3-5).

In general, we find no effects on Math across our various specifications, sample constructions, and bandwidth selections. The one exception is a negative estimate in 2012-13 that is statistically significant at the 9 percent level when including controls for student characteristics, grade fixed effects and baseline mean test score (Table 2A, column 6). This result is not robust to alternate specifications (Table 2A, column 4 and 5) or alternate sample

constructions (Table 6A). The estimated math effect in 2012-13 is also not statistically different from the placebo effect in the Pre-TIF years (Table 7A).

7. TEACHER RETENTION RESULTS

In this section, we examine a potential mechanism for the student achievement impact of TIF: the program's effect on teacher retention. Since inexperienced, less effective teachers are likely to replace teachers that leave high-need schools, a PBCS could improve achievement if it induces more teachers to stay (CITE). Table 8 shows the OLS regression discontinuity estimates of α and γ from Equation 1 (first stage) and Equation 2 (reduced-form) for school exit by a given school year for full-time teachers in TIF-participating districts prior to TIF implementation (2009-10 in Panel A and 2010-11 in Panel B).[31] Standard errors are clustered at the school-level. The reduced-form results in Column 2, Panel B can be interpreted as follows: full-time teachers in a just TIF-eligible school in 2009-10 were 7 pp less likely to leave that school by 2011-12 than full-time teachers in a just ineligible school (in a TIF-participating district) in 2009-10. This estimate of the intent-to-treat effect is significant at the 9 percent level. The estimated effect drops to the 14 percent significance level in 2012-13, but returns to the 8 percent level in 2013-14. Critically, there is no reduction in teacher exit in 2010-11, the TIF planning year, when teachers would likely have made their exit decision prior to being informed about the TIF program (TIF recipients were not announced until the Fall 2010). Panel B shows similar results, but at higher levels of significance, for full-time teachers in 2010-11. Figures 4 and 5 present estimates in results in Panel A and B graphically. As in the case of student achievement, the

[31] Table A3 shows similar results when excluding teachers who ever worked at a school which opened or closed during TIF implementation.

implied LATOT impact on teacher exit is large, 20 percentage points in 2011-12 to 27 percentage points in 2013-14 for 2009-10 full-time teachers.

Table 9 presents the reduced-form estimates from Table 8, but varies the size of the bandwidth around the eligibility threshold (i.e. +/-40, +/-30, and +/-20 percentage points). The coefficient estimates are no longer statistically significant when estimated on the +/-30 and +/-20 windows. This could be a result of insufficient power when using these smaller samples, or it could be that TIF impact estimates in Table 8 are driven by observations far from the eligibility threshold. The latter would be a problem for the internal validity of our RD strategy, as applied to teacher retention.

Table 10 shows the results for the same specification as Table 8, but for year-over-year school exits.[32] Year-over-year exit is an indicator equal to one if a full-time teacher in year *t-1* is not in the same school in year *t*. These results show that the TIF eligibility significantly reduced year-over-year teacher exits in the 2011-12 and 2013-14 (the first and third year of TIF implementation) by 8 percentage points and 6 percentage points (significant only when including covariates), respectively. This retention effect in 2011-12 corresponds to the reading achievement impact in that school year observed in Section 6 (our student achievement data does not contain 2013-14). Figure 7 depicts these results graphically. Table 11 varies the window around the eligibility threshold for these results, using the specification which includes teacher covariates. Only the 2013-14 coefficient estimates remain statistically significant (or nearly significant) for all three window choices. The 2011-12 estimate's significance level drops to 15 percent for the +/-30 window and is not significant for the +/-20 window.

[32] Table A4 shows similar results when excluding teachers who ever worked at a school which opened or closed during TIF implementation.

Table 12 presents tests for discontinuities in teacher covariates at the TIF eligibility threshold. The existence of such discontinuities could indicate that teachers just above and below the threshold are not comparable. This would violate the internal validity of the RD strategy *if* these differences are relevant to the outcome of interest. Each column in the table shows OLS estimates of γ in Equation 1, where the dependent variable is a teacher covariate (years of experience, indicator for master's degree or higher, log of base salary). Each row denotes a different window around the eligibility threshold. We do not detect any discontinuities in the covariates of 2009-10 full-time teachers that are consistent across windows. However, we do see evidence of discontinuities in experience and base salary for 2010-11 full-time teachers. We construct an index of the probability of teacher exit based on teacher covariates and find no evidence of a discontinuity in this index for either 2009-10 or 2010-11 teachers. This suggests that the discontinuities in 2010-11 teacher covariates at the threshold are not meaningful for the outcome of interest and therefore do not violate the identification assumption of the RD strategy.

8. CONCLUSION

We use student and teacher-level data from six TIF-participating districts in the state of Oregon to provide early evidence on how high-profile reforms, implemented as part of Teacher Incentive Fund (TIF), influenced measures of student performance and teacher retention. Leveraging an eligibility threshold in TIF-eligibility based on the percentage of FRL enrollment in a school, we use an RD design to find that Oregon's TIF-reforms *can* produce substantial improvements in reading achievement in selected schools, but we find no improvements in math achievement. We use the same approach to provide suggestive evidence that TIF increased

teacher retention in 2011-12, a possible mechanism for the effect of TIF on student achievement. While significantly different from zero, the imprecision of our estimates of TIF's impact on Reading and retention in 2011-12 results in wide range of possible true effect sizes. Moreover, our estimates provide an upper bound for the impact of TIF if administrators positively selected schools into TIF from the pool of TIF-eligible schools.

Our RD design provides an important complement to the ongoing, federally sponsored randomized trial study in 12 TIF-participating districts. The treatment contrast fielded in the federal evaluation is narrower than that implied by the introduction of PBCS under the TIF program. Specifically, the staff in the experimental control schools receives individualized feedback on their effectiveness, professional-development opportunities, an automatic pay bonus, and additional pay for taking on new roles and responsibilities (e.g., peer mentoring, leading professional development activities). In the treatment schools, staff were eligible for these TIF elements and for differentiated awards based on their performance that were "a little larger on average, than in the control schools" (Max et al. 2014).

The contrast used to identify our RD estimates, leverages more components of the full treatment contrast created by the introduction of a PBCS under TIF because the comparison is between TIF-eligible and ineligible schools. Specifically, teachers in TIF schools experienced more highly trained administrators, higher-quality evaluation feedback, and increased availability of professional development compared to their counterparts in ineligible schools. Furthermore, the compensation contrast in our design compares incentive awards with the status quo, rather than with an automatic pay increase.

Early evidence from the federal evaluation (Max et al. 2014) suggests that implementation fidelity among these awardees has been quite poor. Fewer than half of districts implemented all of the required TIF elements, incentive awards appear to have been insufficiently challenging (most teachers received one), and many teachers may have misunderstood the performance measures and bonuses. In light of these issues, our focus on a single awardee with comparatively high fidelity implementation is likely to provide a more powerful proof of concept than the federal evaluation. Under the relatively favorable conditions in Oregon, we find promising but uneven effects of TIF on achievement.

REFERENCES

Aaronson, D., L Barrow, W Sander (2007). Teachers and Student Achievement in the Chicago Public Schools. Journal of Labor Economics 25(1) 95-135.

Atteberry, A., S. Loeb, and J. Wyckoff. "Do First Impressions Matter? Improvement in EarlyCareer Effectiveness," CALDER Working Paper No. 90, February 2013.

Ballou, D. (2001). "Pay for Performance in Public and Private Schools." Economics of Education Review 20(1): 51-61.

Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications-to-transfer to uncover preferences of teachers and schools. Journal of Policy and Management, 30(1), 88–110.

Camerer, C. and R.M. Hogarth. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," Journal of Risk and Uncertainty 19(1-3), December 1999, 7-42.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood: National Bureau of Economic Research.

Clotfelter, C., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessmentof teacher effectiveness. Journal of Human Resources, 41(4), 778.

Clotfelter, C., E, Glennie, H. Ladd, and J. Vigdor (2008) "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina" Journal of Public Economics, 92: 1352-1370.

Danielson, Charlotte & Thomas L. McGreal. (2000). Teacher evaluation to enhance professional practice. Alexandria, Va.: Association for Supervision and Curriculum Development.

Dee, T.S. and B. Keys (2004). "Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment," Journal of Policy Analysis and Management, 23(3), pages 471-488.

Dee, T.S. and J. Wyckoff (2013). "Incentives, Selection, and Teacher Performance: Evidence from IMPACT," NBER Working Paper 19529.

Fryer, R. (2013). "Teacher Incentives and Student Achievement: Evidence form New York City Public Schools," Journal of Labor Economics 31(2): 373-427.

Fryer, R., S. Levitt, J. List, and S. Sadoff. "Enhancing the Efficacy of Teacher Incentives through Loss Aversion," NBER Working Paper 18237, 2012.

Glazerman, S. and A. Seifullah (2012). An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) After Four Years. Princeton: Mathematica Policy Research.

Glazerman, S., E. Isenberg, S. Dolfin, M. Bleeker, A. Johnson, M. Grider, and M. Jacobus. (2010). Impacts of Comprehensive Teacher Induction: Final Results From a Randomized

Goe, Laura. and Andrew Croft. (2009). Methods of evaluating teacher effectiveness. Washington, DC: National Comprehensive Center for Teacher Quality.

Goldhaber, D., B. Gross and D. Player (2007) “Are Public Schools Really Losing Their “Best”? Assessing the Career Transitions of Teachers and Their Implications for the Quality of the Teacher Workforce,” CALDER working paper.

Hanushek, E. (2007) "The Single Salary Schedule and Other Issues of Teacher Pay" Peabody Journal of Education, 82(4), 574-586.

Hanushek, E., J. Kain, D. O’Brien and S. Rivkin (2005) “The Market for Teacher Quality,” NBER working paper.

Hanushek, E. and S. Rivkin (2012). “The Distribution of Teacher Quality and Implications for Policy,” Annual Review of Economics 4: 131-157.

Imbens, G. and K. Kalyanaraman (2012). “Optimal Bandwidth Choice for the Regression Discontinuity Estimator,” Review of Economic Studies 79 (3): 933-959

Johnson, S. M., and John Papay (2009) Redesigning Teacher Pay, Washington, D.C.: EPI.

Joseph, C. “New Evaluation System for New York Teachers,” New York Times, June 2, 2013, page A20.

Lee, D.S. & Lemieux, T. (2009). Regression discontinuity designs in economics. Journal of Economic Literature, 48, 281-355.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: a density test. Journal of Econometrics, 142(2), 698-714.

McNeil, M. “Feds, States Dicker Over Evaluations,” Education Week, Vol. 32, Issue 12, March 16, 2013, pages 1,24.

Measures of Effective Teaching (MET) Project (2013). Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project’s Three-Year Study.The Bill and Melinda Gates Foundation, January 2013.

Murnane, R. (1984) “Selection and Survival in the Teacher Labor Market,” The Review of Economics and Statistics 66(3) 513-518.

Murnane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. Harvard Educational Review, 56(1), 1–17.

Pianta, R.C. and B.K. Hamre. “Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity,” Educational Researcher, March 2009 vol. 38 no. 2, 109-119.

Rivkin, S., E. Hanushek, and J. Kain (2005). “Teachers, Schools, and Academic Achievement,” Econometrica, 73(2), 417-458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94(2), 247-252.

Rothstein, J. (2012). “Teacher Quality Policy When Supply Matters,” NBER Working Paper No. 18419, September 2012.

Sanders, W.L. and Rivers, J.C. (1996). "Research Project Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," University of Tennessee Value-Added Research and Assessment Center.

Schochet, P. Z., Cook, T., Deke, J., Imbens, G., Lockwood, J.R., Porter, J., & Smith, J. (2010).

Standards for Regression Discontinuity Designs. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf.

Springer, M. (2009) "Rethinking Teacher Compensation Policies: Why Now, Why Again?" in Performance Incentives, M. Springer Ed., Washington DC: Brookings Institution.

Springer, M., D. Ballou, L. Hamilton, V. Le, J.R. Lockwood, D. McCaffrey, M. Pepper, B. Stecher (2010) Teacher Pay for Performance, Experimental Evidence from the Project on Incentives in Teaching, Nashville, TN: National Center on Performance Incentives at Vanderbilt University

Springer, M., J. Pane, V. Le, D. McCaffrey, S. Burns, L. Hamilton, and B. Stecher. “Team Pay for Performance: Experimental Evidence From the Round Rock Pilot Project on Team Incentives,” Educational Evaluation and Policy Analysis December 2012, Vol. 34, No. 4, pp. 367–390.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. Brooklyn, NY: The New Teacher Project.

Weiss, E. and D. Long (2013) Market-Oriented Education Reforms’ Rhetoric Trumps Reality. Broader, Bolder Approach to Education. http://www.epi.org/files/2013/bba-rhetorictrumps-reality.pdf.

Taylor, E.S. and Tyler, J.H. 2012. The effect of evaluation on teacher performance. American Economic Review, 102(7), pp. 3628-3651. 2012.

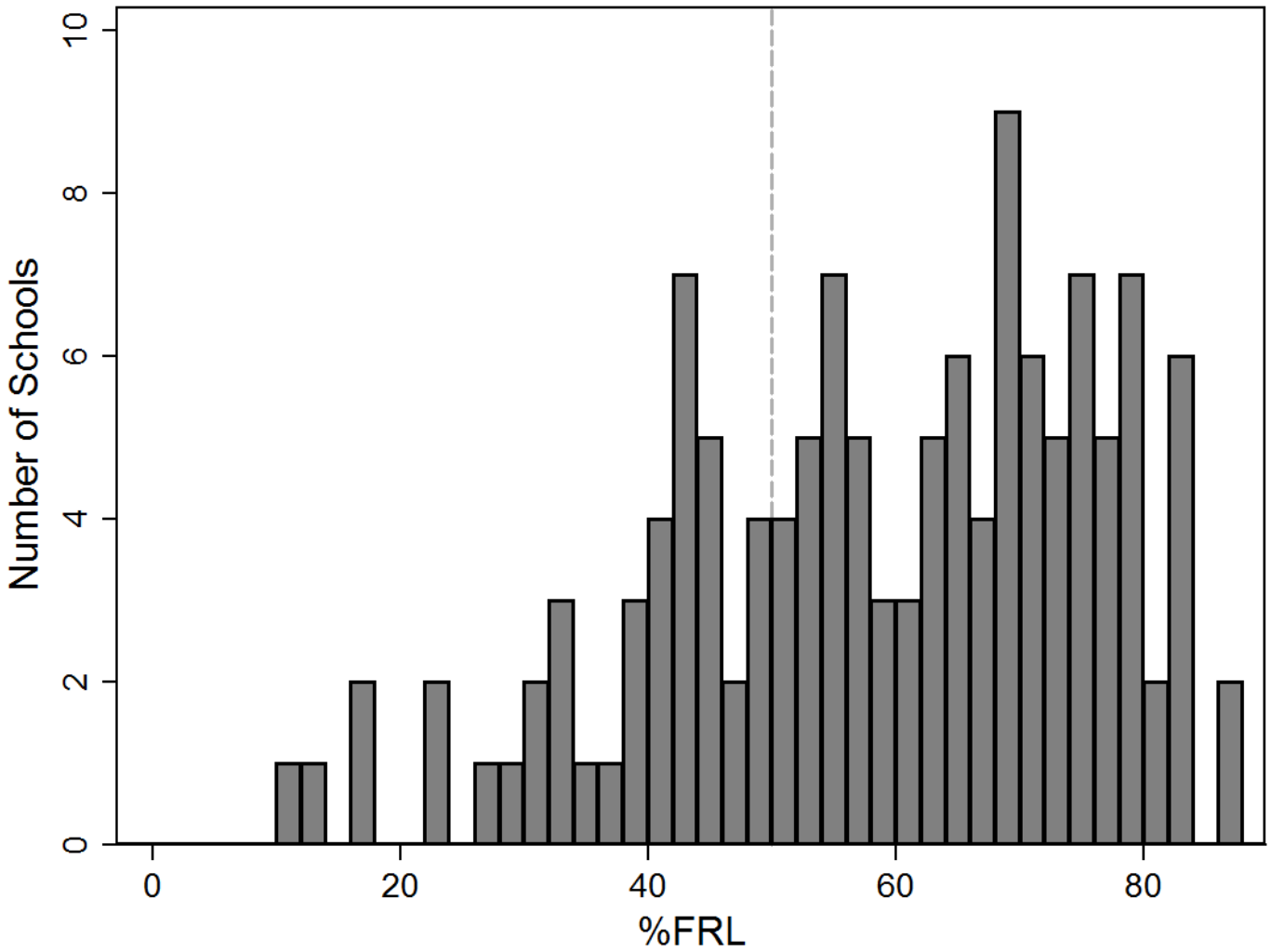
Turque, B. “D.C. teachers likely to ratify contract,” Washington Post, June 1, 2010, B01.

Ujifusa, A. “Buffalo Battle Rages over Evaluations,” Education Week, Vol. 32, Issue 30, May 8, 2013, Page 21.

Yoon , Kwang Suk, Teresa Duncan, Silvia Wen-Yu Lee, Beth Scarloss, and Kathy L. Shapley. 2007. Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement: Issues & Answers Report, REL 2007 ‗ No. 033. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

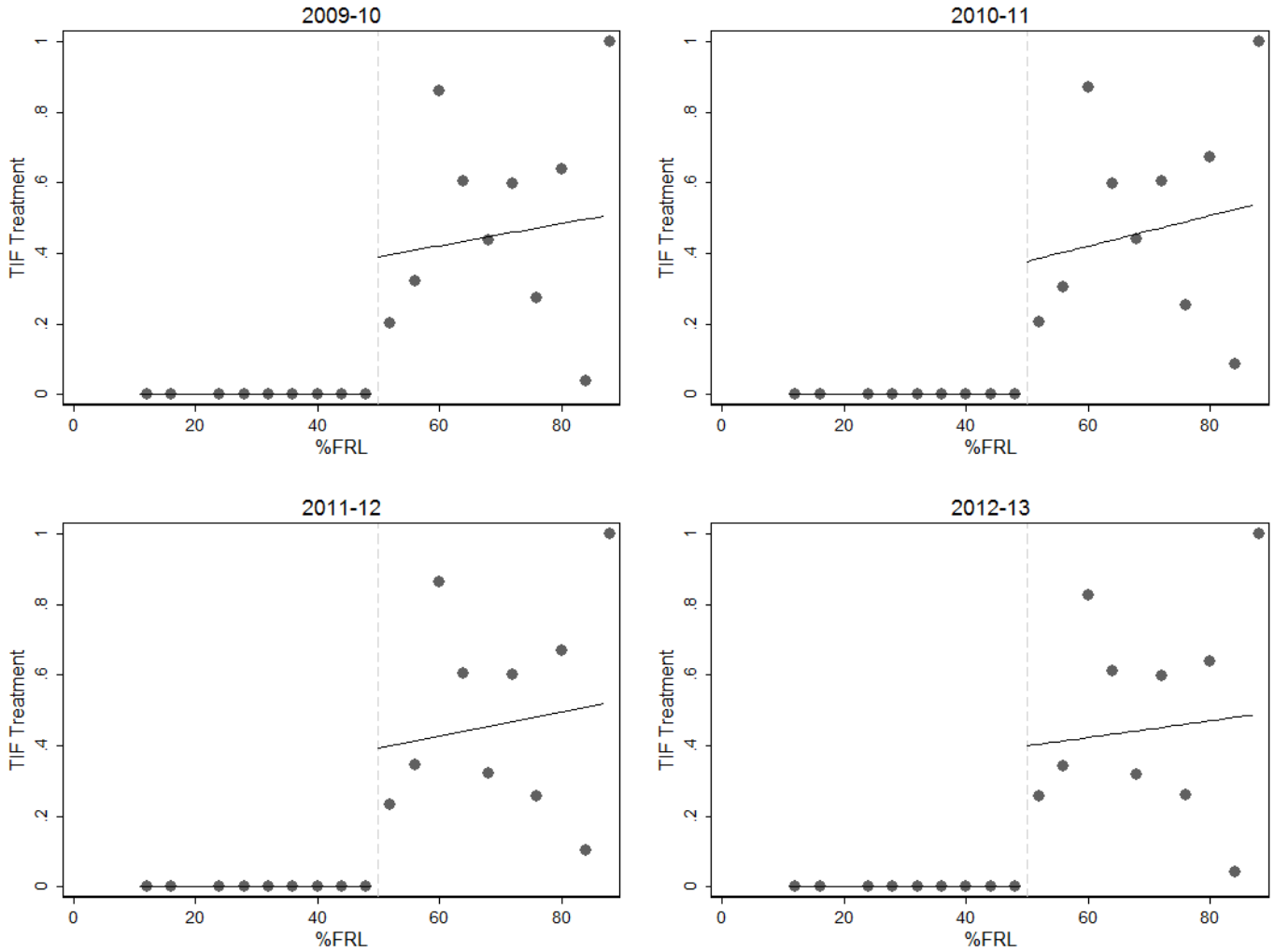
Zubrzycki, Jacklyn. “Districts Abandon Grants Targeting Teacher Quality,” Education Week, Vol. 32, Issue 01, August 22, 2012, pages 1, 20-21.

Figure 1: School-level Distribution of the Assignment Variable (%FRL)



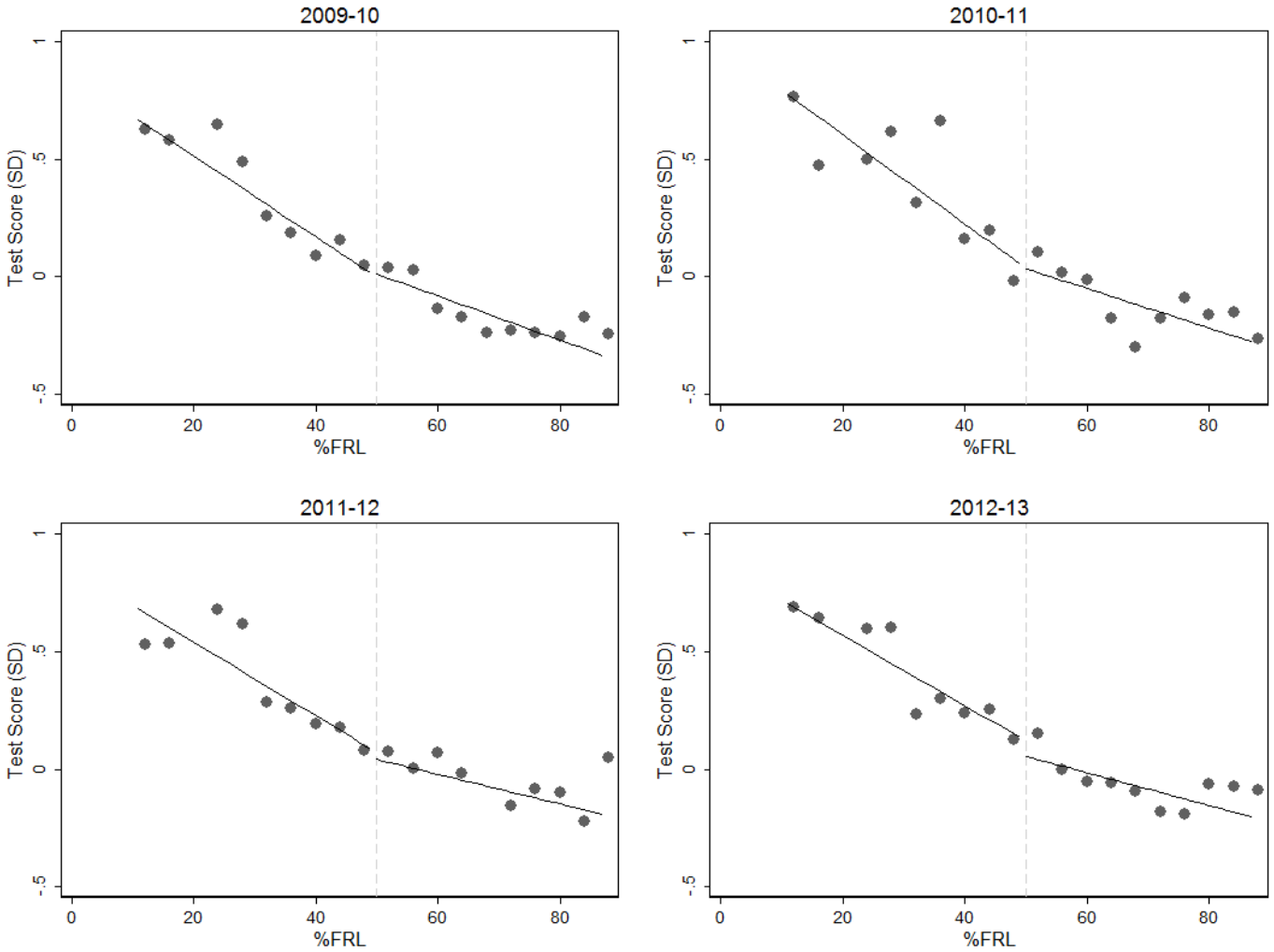
Note: Figure constructed from sample of 2011-12 OAKS reading test-takers in TIF-participating districts (excluding charter schools), collapsed to the school-level and depicted in 2 percentage point %FRL bins. %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 2: First Stage Regression Discontinuity, by School Year



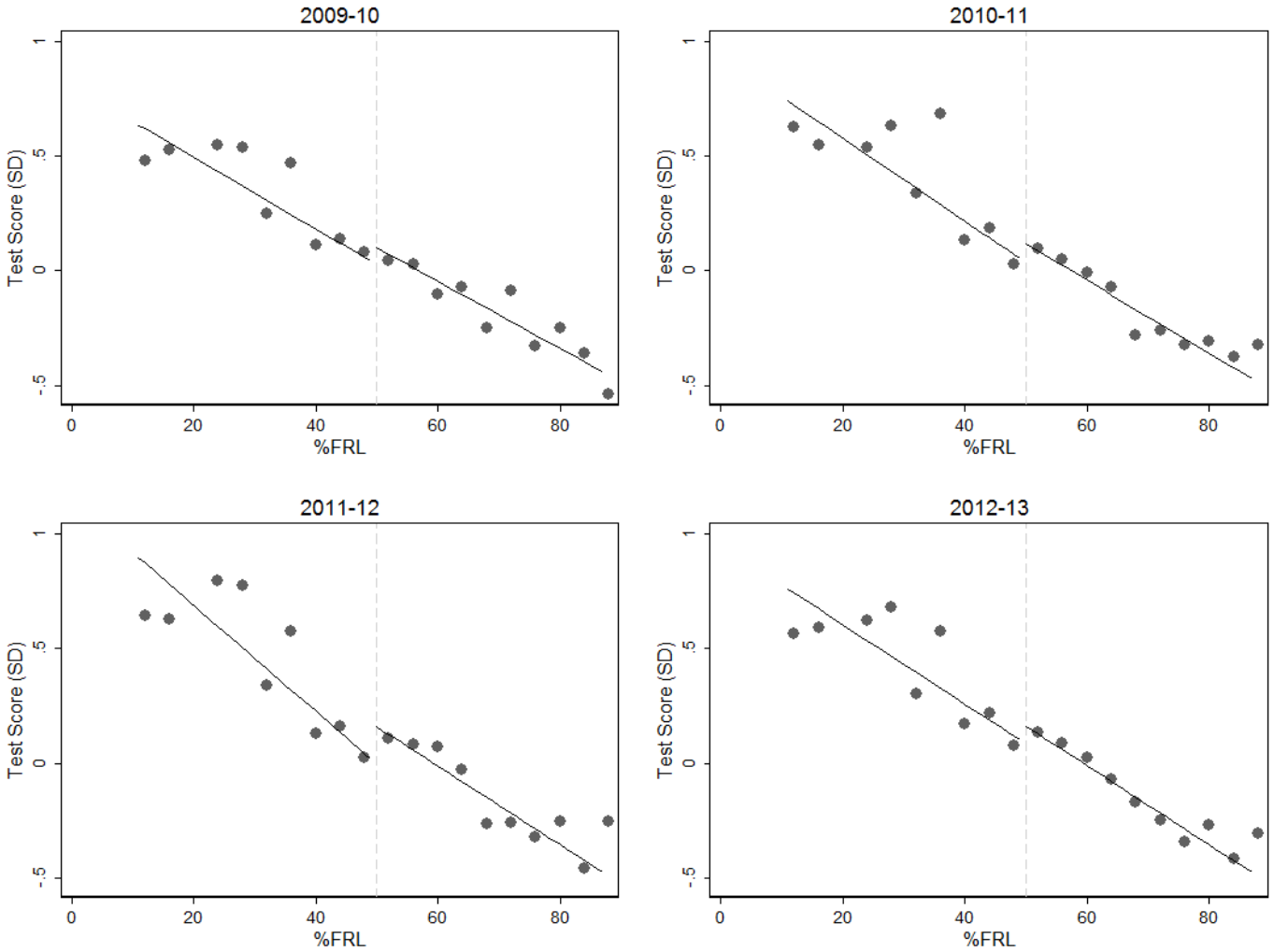
Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean TIF treatment in 4 percentage point %FRL bins. Lines show first stage RD regression (Equation 1). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 3A: Reduced-Form Regression Discontinuity, by School Year (Math)



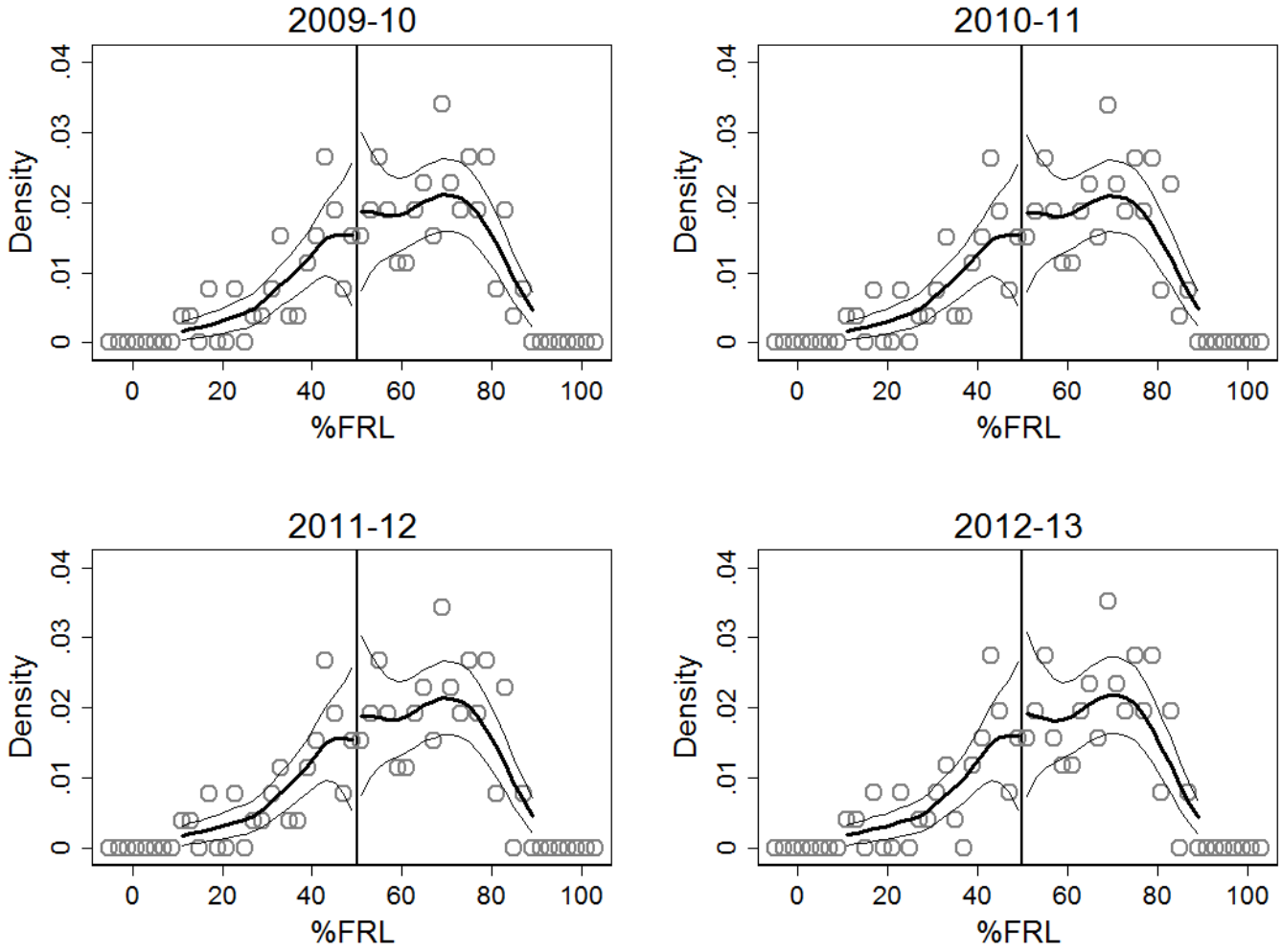
Note: Figure uses the sample of OAKS math test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 3B: Reduced-Form Regression Discontinuity, by School Year (Reading)



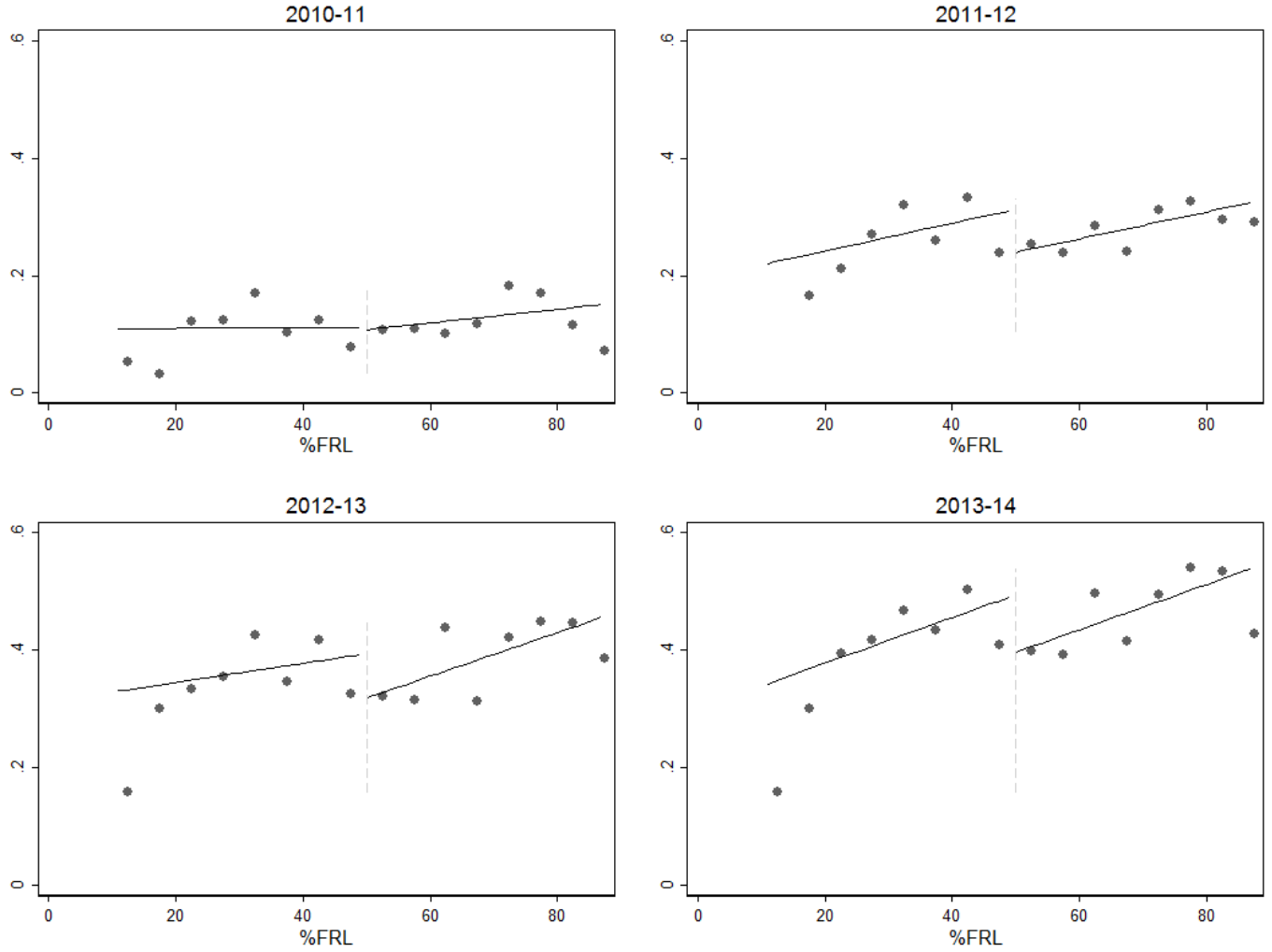
Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 4: School-level McCrary Density Test, by School Year



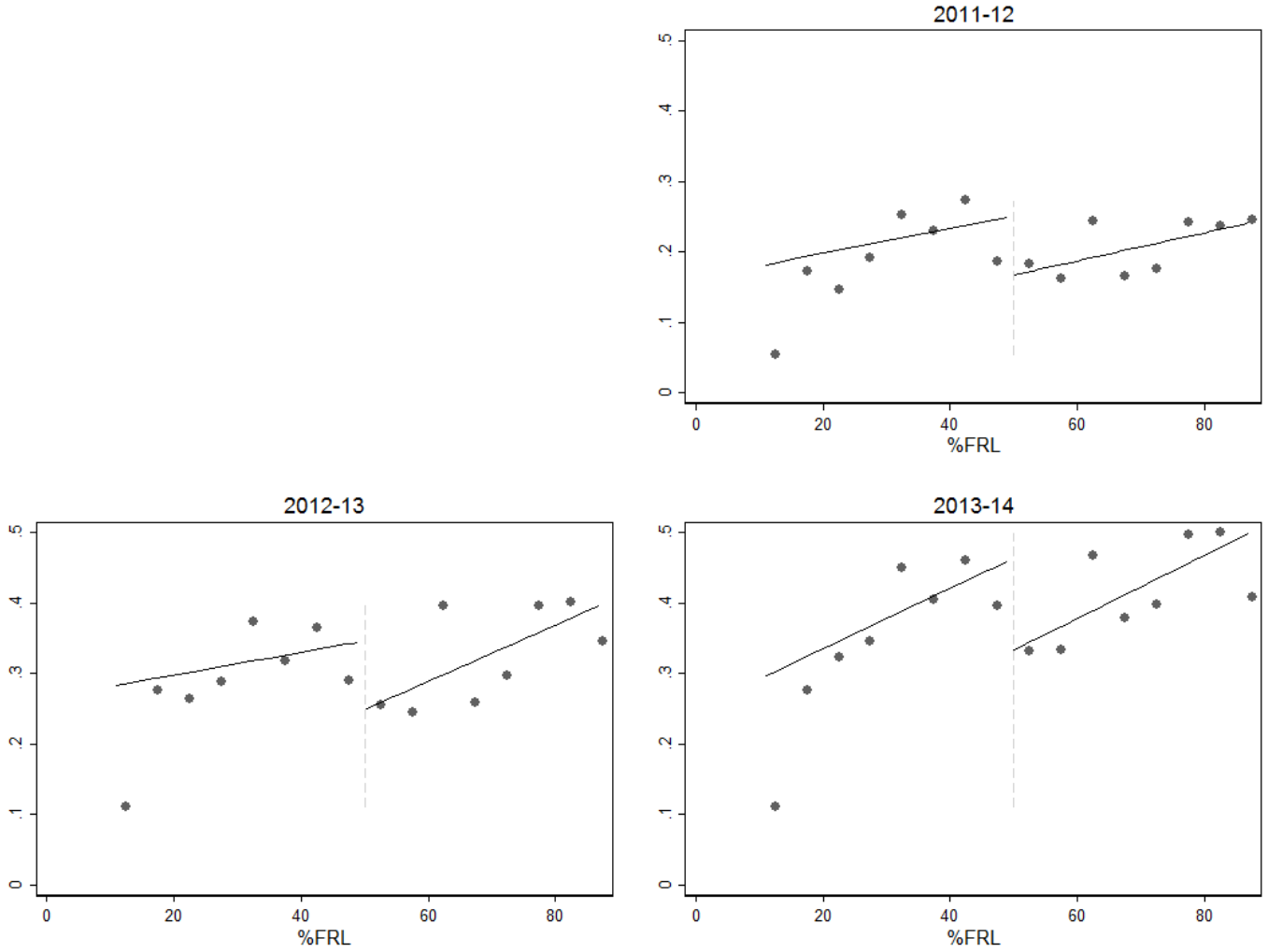
Note: Figure applies the density test from McCrary (2006) to the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools), collapsed to the school-level. Binwidth of 2 is used. %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 5: School Exit by 2009-10 Full-Time Teachers



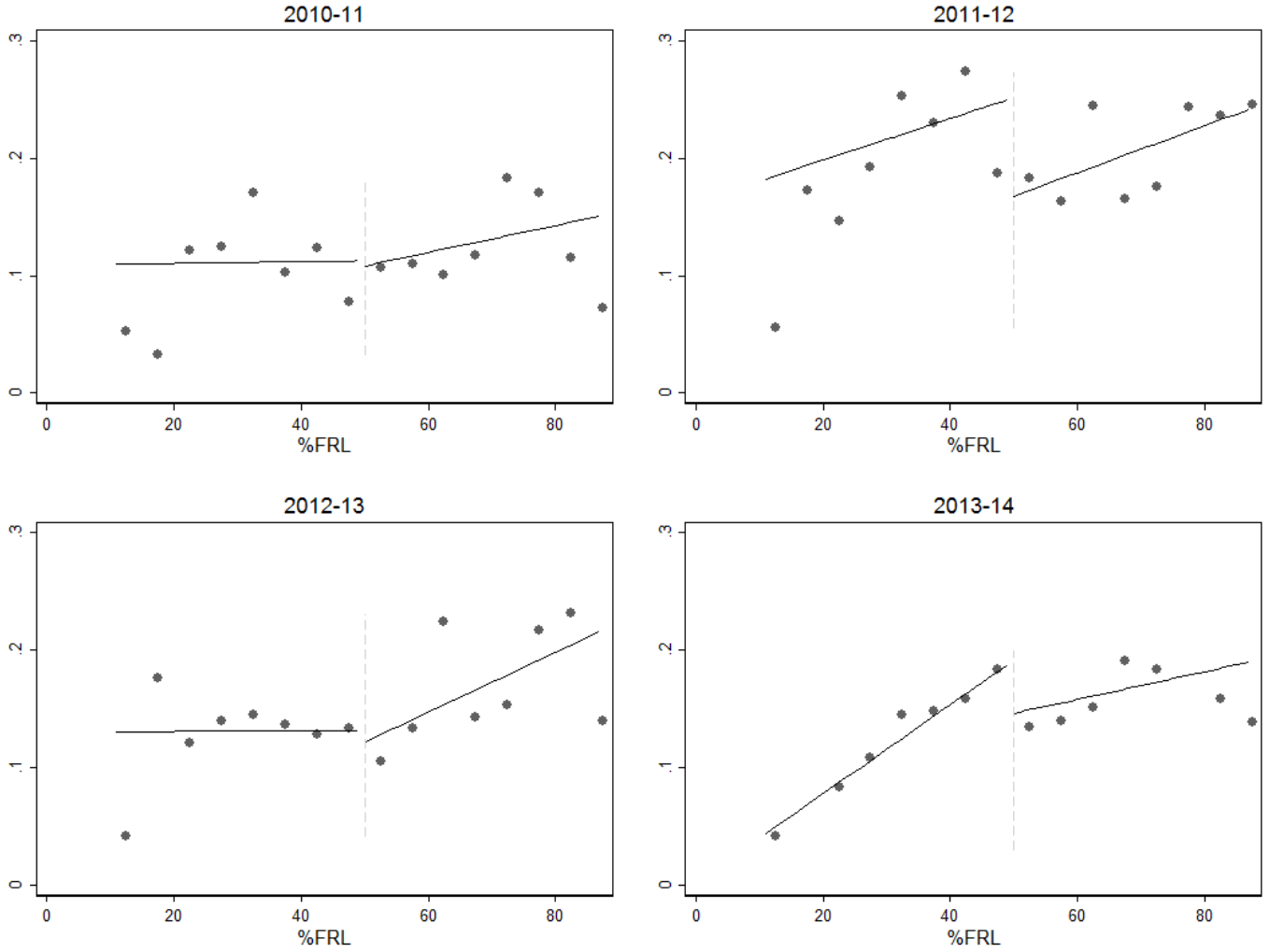
Note: Figure uses the sample of 2009-10 full-time teachers in TIF-participating districts (excluding charter schools). Dots depict school exit likelihoods in 4 percentage point %FRL bins. School exit indicator is equal to one in year t when a teacher that was observed working FT in school s in 2009-10 is no longer observed working FT in school s in year t. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 6: School Exit by 2010-11 Full-Time Teachers



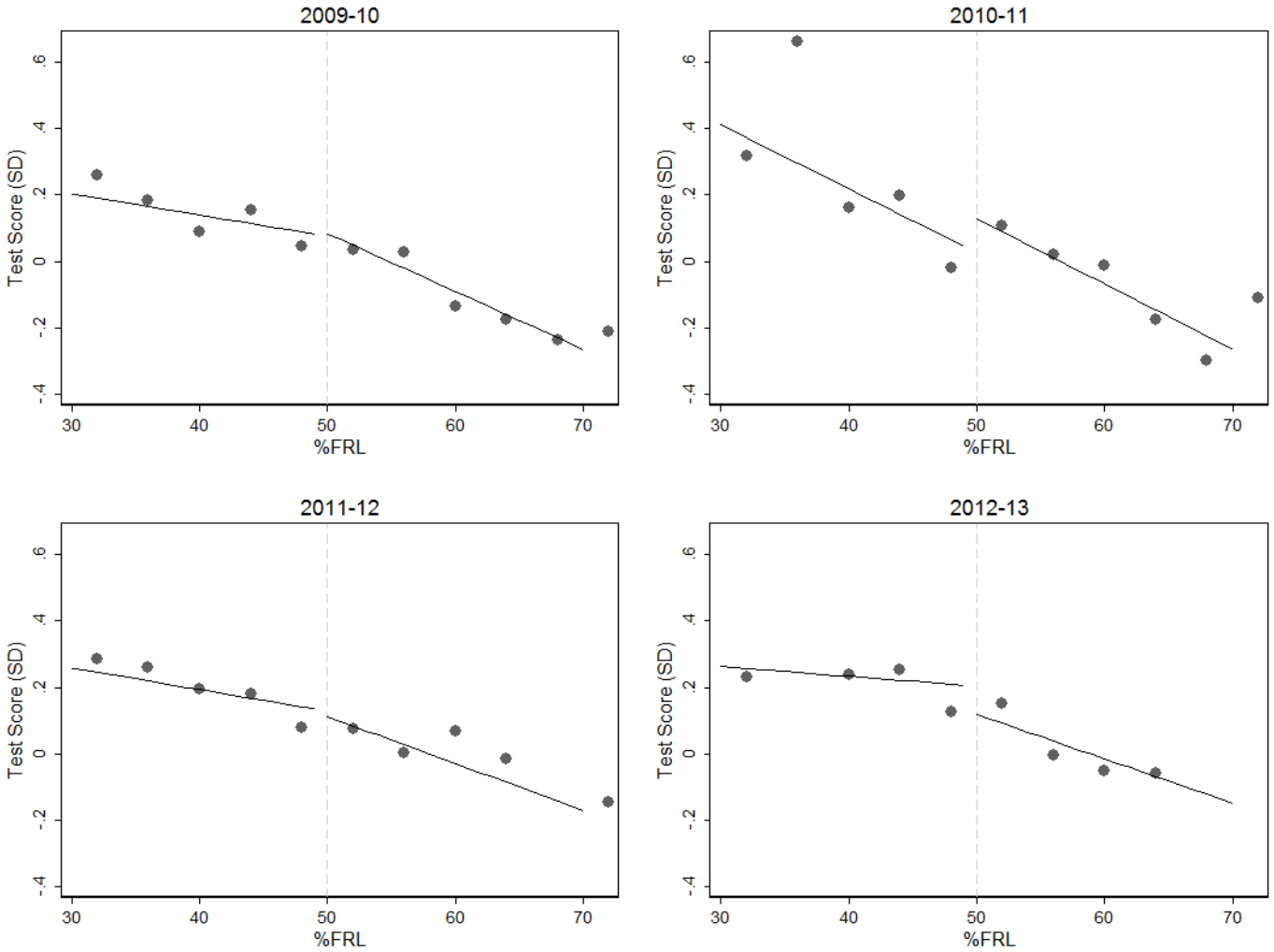
Note: Figure uses the sample of 2010-11 full-time teachers in TIF-participating districts (excluding charter schools). Dots depict school exit likelihoods in 5 percentage point %FRL bins. School exit indicator is equal to one in year t when a teacher that was observed working FT in school s in the base year 2010-11 is no longer observed working FT in school s in year t. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 7: Year-over-Year School Exit by Full-Time Teachers



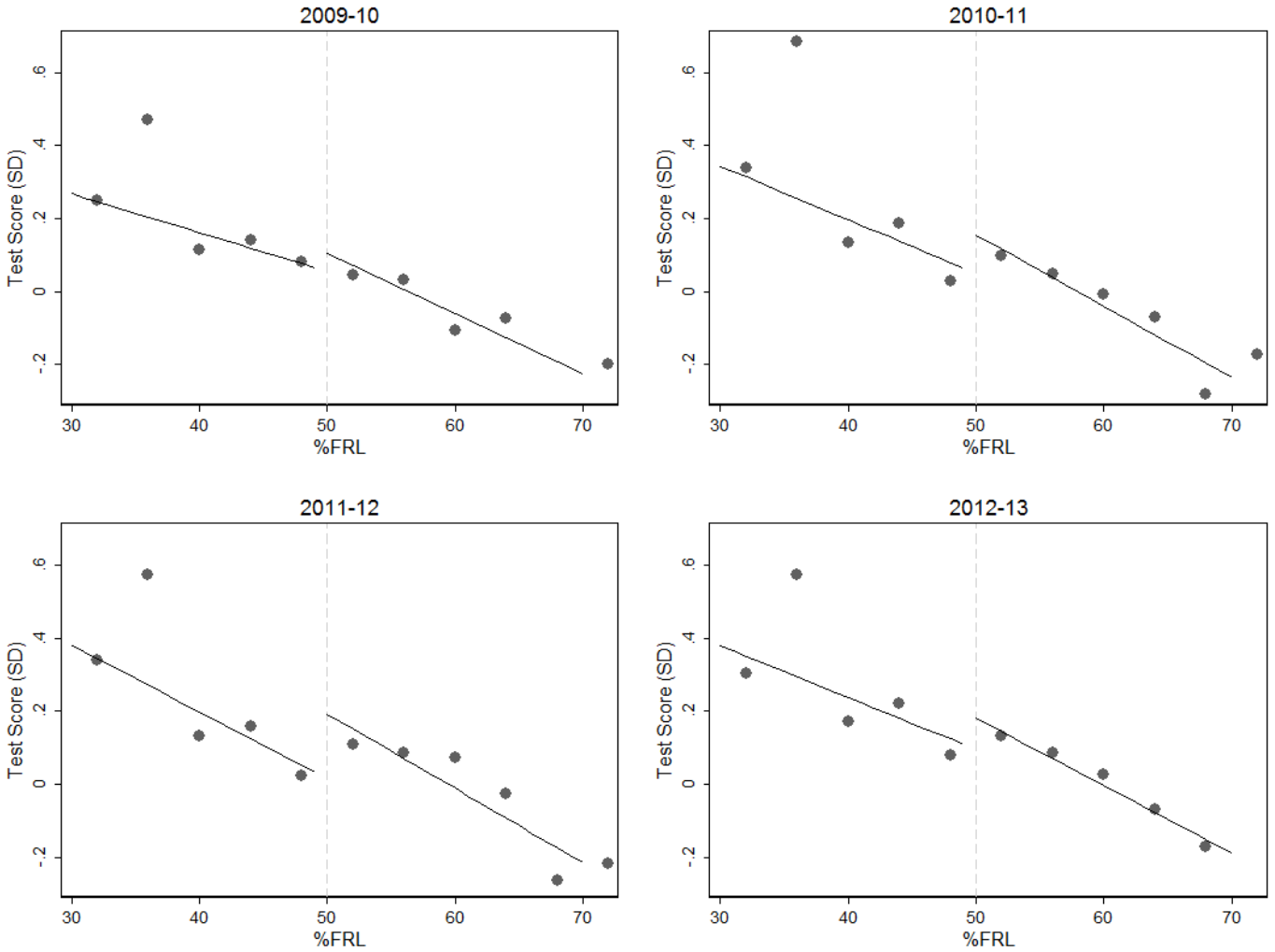
Note: Figure uses the sample of school year t-1 full-time teachers in TIF-participating districts (excluding charter schools). Dots depict school exit likelihoods in 5 percentage point %FRL bins. Year-over-YEar school exit indicator is equal to one in year t when a teacher that was observed working FT in school s in year t-1 is no longer observed working FT in school s in year t. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure A1: Reduced-Form RD for 20-70% FRL Window, by School Year (Math)



Note: Figure uses the sample of OAKS math test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 5 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure A2: Reduced-Form RD for 20-70% FRL Window, by School Year (Reading)



Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Table 1 - Summary Statistics for TIF-Participating Districts (2009-10)

	All Schools (1)	TIF-Ineligible Schools (2)	TIF-Eligible Non-TIF Schools (3)	TIF Schools* (4)
Sample (Counts)				
Test-takers	44,624	13,935	17,146	13,660
Teachers	3,270	974	1,313	983
Schools	132	41	51	40
Student Achievement				
Math SD Score	-0.03	0.19	-0.10	-0.19
Reading SD Score	-0.03	0.21	-0.11	-0.19
Student Covariates				
Economically Disadvantaged	0.59	0.36	0.69	0.68
White	0.67	0.79	0.60	0.65
Hispanic	0.25	0.13	0.32	0.28
Black	0.01	0.01	0.01	0.01
Special Education	0.17	0.16	0.17	0.17
Grade 3	0.13	0.15	0.12	0.12
Grade 4	0.14	0.16	0.13	0.12
Grade 5	0.13	0.16	0.12	0.13
Grade 6	0.13	0.13	0.14	0.13
Grade 7	0.13	0.13	0.14	0.12
Grade 8	0.13	0.12	0.14	0.12
High School	0.21	0.15	0.21	0.27
Teacher Covariates				
Years of Experience	12.33	13.22	12.27	11.54
MA Degree or Higher	0.63	0.61	0.63	0.63
Base Salary	52,102	52,747	52,197	51,336
Elementary School	0.46	0.51	0.44	0.43

Each entry is the mean for the given variable (unless otherwise noted).

Includes data on all OAKS test-takers for 2009-2010 in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer).

Sample counts and student covariate means are for reading test-takers (similar for math test-takers).

Table 2A - RD Estimates of TIF Impact on Math Achievement (SD Scores)

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage	0.38 ***	0.39 ***	0.39 ***	0.43 ***	0.43 ***	0.44 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	-0.03	-0.05	-0.04	-0.07	-0.08	-0.09 *
	(0.09)	(0.08)	(0.05)	(0.08)	(0.06)	(0.05)
First Stage F	*9.04*	*9.10*	*9.12*	*10.98*	*11.05*	*11.04*
Obs	42,816	*41,710*	*41,341*	*41,989*	*40,917*	*40,530*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Test Score Mean 2009-10 (Grade x School)			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table 2B - RD Estimates of TIF Impact on Reading Achievement (SD Scores)

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage	0.39 ***	0.39 ***	0.39 ***	0.40 ***	0.40 ***	0.41 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.16 **	0.12 **	0.10 **	0.08	0.06	0.05
	(0.07)	(0.06)	(0.05)	(0.07)	(0.05)	(0.05)
First Stage F	*9.15*	*9.12*	*9.36*	*9.64*	*9.73*	*9.89*
Obs	42,730	*41,567*	*41,199*	*41,974*	*40,853*	*40,466*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Test Score Mean 2009-10 (Grade x School)			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Reduced Form regressions: y = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Standard errors clustered at the school level are in parentheses.

Table 3A - RD Estimates of TIF Impact on Math Achievement (SD Scores) for Various Bandwidths

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
FULL SAMPLE						
First Stage	0.38 ***	0.39 ***	0.39 ***	0.43 ***	0.43 ***	0.44 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	-0.03	-0.05	-0.04	-0.07	-0.08	-0.09 *
	(0.09)	(0.08)	(0.05)	(0.08)	(0.06)	(0.05)
Obs	*42,816*	*41,710*	*41,341*	*41,989*	*40,917*	*40,530*
20 < FRL < 80						
First Stage	0.37 ***	0.37 ***	0.38 ***	0.41 ***	0.41 ***	0.42 ***
	(0.13)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Reduced Form (Elgibility)	0.01	-0.01	-0.03	-0.04	-0.07	-0.10 *
	(0.10)	(0.09)	(0.06)	(0.09)	(0.07)	(0.05)
Obs	*38,631*	*37,695*	*37,414*	*37,950*	*37,045*	*36,748*
30 < FRL < 70						
First Stage	0.35 **	0.35 **	0.35 **	0.40 **	0.40 **	0.40 **
	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)	(0.16)
Reduced Form (Elgibility)	-0.02	-0.02	-0.02	-0.09	-0.09	-0.11 **
	(0.12)	(0.10)	(0.06)	(0.10)	(0.07)	(0.05)
Obs	*29,494*	*28,885*	*28,658*	*28,771*	*28,172*	*27,938*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Test Score Mean 2009-10 (Grade x School)			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table 3B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) for Various Bandwidths

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
FULL SAMPLE						
First Stage	0.39 ***	0.39 ***	0.39 ***	0.40 ***	0.40 ***	0.41 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.16 **	0.12 **	0.10 **	0.08	0.06	0.05
	(0.07)	(0.06)	(0.05)	(0.07)	(0.05)	(0.05)
Obs	*42,730*	*41,567*	*41,199*	*41,974*	*40,853*	*40,466*
20 < FRL < 80						
First Stage	0.39 ***	0.38 ***	0.38 ***	0.39 ***	0.39 ***	0.39 ***
	(0.14)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)
Reduced Form (Elgibility)	0.24 ***	0.21 ***	0.18 ***	0.14 *	0.11 *	0.09
	(0.08)	(0.07)	(0.06)	(0.08)	(0.06)	(0.05)
Obs	*38,462*	*37,500*	*37,220*	*37,758*	*36,832*	*36,536*
30 < FRL < 70						
First Stage	0.38 **	0.37 **	0.36 **	0.39 **	0.39 **	0.39 **
	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
Reduced Form (Elgibility)	0.17 *	0.17 **	0.13 **	0.09	0.08	0.04
	(0.09)	(0.08)	(0.05)	(0.09)	(0.07)	(0.05)
Obs	*29,389*	*28,752*	*28,526*	*28,626*	*28,009*	*27,774*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Test Score Mean 2009-10 (Grade x School)			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table 4A - RD Estimates of TIF Impact on Math Achievement (SD Scores) in Pre-Treatment Period

	TIF Planning Year								
	2007-08		2008-09		2009-10		2010-11		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First Stage	0.36 ***	0.35 ***	0.38 ***	0.37 ***	0.36 ***	0.36 ***	0.39 ***	0.38 ***	0.39 ***
	(0.12)	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.01	-0.02	0.09	0.05	0.00	-0.02	-0.01	-0.03	-0.04
	(0.08)	(0.07)	(0.07)	(0.06)	(0.08)	(0.07)	(0.07)	(0.07)	(0.05)
First Stage F	*8.73*	*8.44*	*9.43*	*8.90*	*8.64*	*8.36*	*9.13*	*9.02*	*9.18*
Obs	*43,571*	*43,571*	*44,920*	*44,920*	*44,832*	*44,830*	*45,570*	*44,438*	*44,180*
Specifications									
Student Characteristics		X		X		X		X	X
Grade Fixed Effects		X		X		X		X	X
Test Score Mean 2009-10 (Grade x School)									X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table 4B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) in Pre-Treatment Period

	TIF Planning Year								
	2007-08		2008-09		2009-10		2010-11		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First Stage	0.40 ***	0.39 ***	0.42 ***	0.41 ***	0.39 ***	0.38 ***	0.38 ***	0.38 ***	0.38 ***
	(0.12)	(0.12)	(0.13)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.05	0.01	0.04	0.00	0.07	0.02	0.08	0.04	0.03
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.05)
First Stage F	*10.30*	*9.76*	*10.96*	*10.56*	*9.83*	*9.33*	*9.14*	*8.90*	*9.16*
Obs	*41,070*	*41,070*	*43,518*	*43,517*	*43,431*	*43,431*	*44,639*	*43,526*	*43,266*
Specifications									
Student Characteristics		X		X		X		X	X
Grade Fixed Effects		X		X		X		X	X
Test Score Mean 2009-10 (Grade x School)									X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Reduced Form regressions: y = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Standard errors clustered at the school level are in parentheses.

Table 5 - RD Estimates of TIF Impact on Test-Taker Covariates for Various Bandwidths

	2011-12					2012-13				
				Predicted Score Index					Predicted Score Index	
	Econ Disadv	White	Special Ed	Math	Reading	Econ Disadv	White	Special Ed	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FULL SAMPLE										
Reduced Form (Elgibility)	-0.01	0.12 ***	0.01	0.02	0.04	0.00	0.10 ***	0.02 *	0.02	0.02
	(0.03)	(0.04)	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)	(0.03)	(0.02)
Obs	*42,730*	*41,567*	*42,730*	*41,710*	*41,567*	*41,974*	*40,853*	*41,974*	*40,917*	*40,853*
20 < FRL < 80										
Reduced Form (Elgibility)	-0.04	0.11 ***	0.02	0.03	0.04	-0.05	0.11 ***	0.02 *	0.04	0.04
	(0.03)	(0.04)	(0.01)	(0.03)	(0.03)	(0.03)	(0.04)	(0.01)	(0.03)	(0.03)
Obs	*38,462*	*37,500*	*38,462*	*37,695*	*37,500*	*37,758*	*36,832*	*37,758*	*37,045*	*36,832*
30 < FRL < 70										
Reduced Form (Elgibility)	-0.03	0.05	0.02	-0.01	0.01	-0.03	0.05	0.02	0.01	0.01
	(0.04)	(0.05)	(0.01)	(0.04)	(0.03)	(0.04)	(0.05)	(0.01)	(0.04)	(0.03)
Obs	*29,389*	*28,752*	*29,389*	*28,885*	*28,752*	*28,626*	*28,009*	*28,626*	*28,172*	*28,009*

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS reading test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Columns 1, 2, 3 and 6,7,8 are estimated using OAKS reading test-takers (results are similar for OAKS math test-takers).

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table 6A - RD Estimates of TIF Impact on Math Achievement (SD Scores) by Cohort

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
Grade 3-5 Students in 2009-2010						
First Stage	0.46 ***	0.47 ***	0.47 ***			
	(0.15)	(0.15)	(0.15)			
Reduced Form (Elgibility)	0.04	0.01	-0.10			
	(0.08)	(0.07)	(0.06)			
First Stage F	*9.95*	*10.05*	*10.38*			
Obs	19,104	*18,605*	*18,605*			
Grade 3-6 Students in 2010-2011						
First Stage	0.46 ***	0.46 ***	0.46 ***	0.47 ***	0.47 ***	0.47 ***
	(0.13)	(0.13)	(0.13)	(0.15)	(0.15)	(0.15)
Reduced Form (Elgibility)	0.07	0.02	0.02	0.00	-0.04	-0.07
	(0.08)	(0.08)	(0.06)	(0.08)	(0.07)	(0.07)
First Stage F	*13.14*	*13.45*	*13.48*	*9.78*	*9.89*	*9.99*
Obs	20,155	*19,651*	*19,651*	*18,476*	*18,019*	*18,019*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Individual Student's Baseline Test Score			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Reduced Form regressions: y = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Standard errors clustered at the school level are in parentheses.

Table 6B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) by Cohort

	2011-12			2012-13		
	(1)	(2)	(3)	(4)	(5)	(6)
Grade 3-6 Students in 2009-2010						
First Stage	0.46 ***	0.47 ***	0.47 ***			
	(0.15)	(0.15)	(0.15)			
Reduced Form (Elgibility)	0.25 ***	0.19 ***	0.13 **			
	(0.07)	(0.07)	(0.06)			
First Stage F	*9.89*	*9.98*	*10.16*			
Obs	18,980	*18,487*	*18,487*			
Grade 3-6 Students in 2010-2011						
First Stage	0.46 ***	0.46 ***	0.46 ***	0.46 ***	0.47 ***	0.47 ***
	(0.13)	(0.13)	(0.13)	(0.15)	(0.15)	(0.15)
Reduced Form (Elgibility)	0.27 ***	0.19 ***	0.10 ***	0.20 ***	0.12 *	0.00
	(0.07)	(0.07)	(0.03)	(0.08)	(0.07)	(0.04)
First Stage F	*13.06*	*13.37*	*13.47*	*9.70*	*9.79*	*10.00*
Obs	20,003	*19,504*	*19,504*	*18,347*	*17,892*	*17,892*
Specifications						
Student Characteristics		X	X		X	X
Grade Fixed Effects		X	X		X	X
Individual Student's Baseline Test Score			X			X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Reduced Form regressions: y = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

Standard errors clustered at the school level are in parentheses.

Table 7A - Difference-in-RD Reduced-Form Estimates of TIF Impact on Math Achievement (SD Scores)

	2008-09 to 2012-13		2011-12 and 2012-13		
	(1)	(2)	(3)	(4)	(5)
Eligibility	0.04	0.01	-0.03	-0.05	-0.05
	(0.04)	(0.04)	(0.09)	(0.08)	(0.05)
Eligibility X (SY 2010-2011)	-0.04	-0.05			
	(0.09)	(0.08)			
Eligibility X (SY 2011-2012)	-0.07	-0.06			
	(0.10)	(0.08)			
Eligibility X (SY 2012-2013)	-0.10	-0.09	-0.03	-0.03	-0.04
	(0.09)	(0.07)	(0.12)	(0.10)	(0.07)
Obs	*263,698*	*260,386*	*84,805*	*82,627*	*81,871*
Specifications					
Student Characteristics		X		X	X
Grade Fixed Effects		X		X	X
Test Score Mean 2009-10 (Grade x School)					X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers 2008-2009 to 2012-2013 in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Pre-TIF years (2008-2009 to 2009-2010) are grouped together.

Standard errors clustered at the school x year level are in parentheses.

Table 7B - Difference-in-RD Reduced-Form Estimates of TIF Impact on Reading Achievement (SD Scores)

	2008-09 to 2012-13		2011-12 and 2012-13		
	(1)	(2)	(3)	(4)	(5)
Eligibility	0.05	0.01	0.16 **	0.12 **	0.10 *
	(0.03)	(0.03)	(0.07)	(0.06)	(0.05)
Eligibility X (SY 2010-2011)	0.03	0.04			
	(0.07)	(0.07)			
Eligibility X (SY 2011-2012)	0.11	0.12 *			
	(0.08)	(0.07)			
Eligibility X (SY 2012-2013)	0.03	0.06	-0.08	-0.06	-0.06
	(0.08)	(0.06)	(0.10)	(0.08)	(0.07)
Obs	*257,362*	*253,964*	*84,704*	*82,420*	*81,665*
Specifications					
Student Characteristics		X		X	X
Grade Fixed Effects		X		X	X
Test Score Mean 2009-10 (Grade x School)					X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers 2008-2009 to 2012-2013 in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Pre-TIF years (2008-2009 to 2009-2010) are grouped together.

Standard errors clustered at the school x year level are in parentheses.

Table 8 - RD Estimates of TIF Impact on School Exit Likelihood of Teachers

	School Year							
	2010-11		2011-12		2012-13		2013-14	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) 2009-10 Full-Time Teachers								
Reduced Form	0.00	0.00	-0.07 *	-0.08 *	-0.08	-0.08	-0.10 *	-0.10 *
	(0.02)	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)
First Stage	0.36 ***	0.36 ***	0.36 ***	0.36 ***	0.36 ***	0.36 ***	0.36 ***	0.36 ***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate	0.12	0.12	0.28	0.28	0.38	0.38	0.46	0.46
First Stage F	*8.78*	*8.79*	*8.78*	*8.79*	*8.78*	*8.79*	*8.78*	*8.79*
Obs	*3,270*	*3,270*	*3,270*	*3,270*	*3,270*	*3,270*	*3,270*	*3,270*
B) 2010-11 Full-Time Teachers								
Reduced Form			-0.08 **	-0.08 **	-0.10 *	-0.09 *	-0.13 **	-0.12 **
			(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)
First Stage			0.34 ***	0.34 ***	0.34 ***	0.34 ***	0.34 ***	0.34 ***
			(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate			0.21	0.21	0.32	0.32	0.41	0.41
First Stage F			*7.78*	*7.65*	*7.78*	*7.65*	*7.78*	*7.65*
Obs			*3,296*	*3,296*	*3,296*	*3,296*	*3,296*	*3,296*
Specification								
Teacher Covariates		*X*		*X*		*X*		*X*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification).

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

First Stage regressions: $TIF = p + q*TIF + r*z*(z>=0) + s*z*(z<0)$

Data includes all base year (2009-10 or 2010-11) full-time (FTE>0.9) teachers in non-charter schools in participating districts with FRL data.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

School Exit indicator is equal to one in year t when a teacher that was observed working FT in school s in the base year (2009-10 or 2010-11) is no longer observed working FT in school s in year t.

Teacher covariates include: cubic in experience, log base salary, indicator for MA or above, indicator for Special Ed, indicator for Elementary School teacher.

Standard errors clustered at school-level are in parentheses.

Table 9 - RD Estimates of TIF Impact on School Exit Likelihood of Teachers -Various Windows

	2009-10 Full-Time Teachers				2010-11 Full-Time Teachers		
	2010-2011	2011-2012	2012-2013	2013-2014	2011-2012	2012-2013	2013-2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FRL[10,90]							
Reduced Form	0.00	-0.08 *	-0.08	-0.10 *	-0.08 **	-0.09 *	-0.12 **
	(0.02)	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)	(0.06)
First Stage F	*8.79*	*8.79*	*8.79*	*8.79*	*7.65*	*7.65*	*7.65*
Obs	*3,270*	*3,270*	*3,270*	*3,270*	*3,296*	*3,296*	*3,296*
FRL[20,80]							
Reduced Form	0.00	-0.06	-0.06	-0.08	-0.06	-0.07	-0.11 *
	(0.03)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.06)
First Stage F	*7.43*	*7.43*	*7.43*	*7.43*	*6.79*	*6.79*	*6.79*
Obs	*2,928*	*2,928*	*2,928*	*2,928*	*2,935*	*2,935*	*2,935*
FRL[30,70]							
Reduced Form	0.02	-0.02	-0.02	-0.05	-0.02	-0.02	-0.06
	(0.03)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
First Stage F	*4.32*	*4.32*	*4.32*	*4.32*	*4.55*	*4.55*	*4.55*
Obs	*2,243*	*2,243*	*2,243*	*2,243*	*2,272*	*2,272*	*2,272*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification with teacher covariates). All schools in participating districts with FRL data.

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at school-level are in parentheses.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

Table 10 - RD Estimates of TIF Impact on Year-over-Year School Exit Likelihood of Teachers

	School Year									
	2009-10		2010-11		2011-12		2012-13		2013-14	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reduced Form	0.06	0.06	0.00	0.00	-0.08 **	-0.08 **	-0.01	-0.01	-0.05	-0.06 *
	(0.04)	(0.04)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
First Stage	0.37 ***	0.37 ***	0.36 ***	0.36 ***	0.34 ***	0.34 ***	0.34 ***	0.33 ***	0.34 ***	0.34 ***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate	0.16	0.16	0.12	0.12	0.21	0.21	0.16	0.16	0.16	0.16
First Stage F	*8.96*	*8.95*	*8.78*	*8.79*	*7.78*	*7.65*	*7.93*	*7.76*	*8.03*	*7.96*
Obs	*130*	*130*	*131*	*131*	*133*	*133*	*131*	*131*	*128*	*128*
Specification										
Teacher Covariates		*X*		*X*		*X*		*X*		*X*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification).

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

First Stage regressions: $TIF = p + q*TIF + r*z*(z>=0) + s*z*(z<0)$

Data includes all base year (2009-10 or 2010-11) full-time (FTE>0.9) teachers in non-charter schools in participating districts with FRL data.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

School Exit indicator is equal to one in year t when a teacher that was observed working FT in school s in the prior year is no longer observed working FT in school s in year t.

Teacher covariates include: cubic in experience, log base salary, indicator for MA or above, indicator for Special Ed, indicator for Elementary School teacher.

Standard errors clustered at school-level are in parentheses.

Table 11 - RD Estimates of TIF Impact on Year-over-Year School Exit Likelihood of Teachers- Various Windows

	School Year				
	2009-10	2010-2011	2011-2012	2012-2013	2013-2014
	(1)	(2)	(3)	(4)	(5)
FRL[10,90]					
Reduced Form	0.06	0.00	-0.08 **	-0.01	-0.06 *
	(0.04)	(0.02)	(0.04)	(0.04)	(0.03)
First Stage F	*8.95*	*8.79*	*7.65*	*7.76*	*7.96*
Obs	*3,452*	*3,270*	*3,296*	*3,024*	*3,001*
FRL[20,80]					
Reduced Form	-0.01	0.00	-0.06	-0.01	-0.07 *
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
First Stage F	*7.78*	*7.43*	*6.79*	*7.19*	*7.16*
Obs	*3,052*	*2,928*	*2,935*	*2,689*	*2,671*
FRL[30,70]					
Reduced Form	-0.02	0.02	-0.02	0.00	-0.07
	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)
First Stage F	*4.39*	*4.32*	*4.55*	*4.68*	*4.57*
Obs	*2,355*	*2,243*	*2,272*	*2,081*	*2,038*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification with teacher covariates).

All schools in participating districts with FRL data.

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at school-level are in parentheses.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

Table 12 - RD Estimates of Teacher Covariate Balance

	2009-10 Full-Time Teachers				2010-11 Full-Time Teachers			
	Experience	MA+	Log Base Salary	Exit Index	Experience	MA+	Log Base Salary	Exit Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FRL[0,100]								
Reduced Form	0.97	-0.03	0.01	0.00	1.74 ***	-0.01	0.03 **	0.00
	(0.64)	(0.06)	(0.02)	(0.00)	(0.57)	(0.05)	(0.01)	(0.01)
Exit Rate	12.33	0.63	10.84	0.12	12.54	0.67	10.85	0.21
First Stage F	*8.78*	*8.78*	*8.78*	*8.78*	*7.78*	*7.78*	*7.78*	*7.78*
Obs	*3,270*	*3,270*	*3,270*	*3,270*	*3,296*	*3,296*	*3,296*	*3,296*
FRL[20,80]								
Reduced Form	1.37 *	0.02	0.02	0.00	1.90 ***	0.02	0.04 **	0.00
	(0.71)	(0.06)	(0.02)	(0.00)	(0.62)	(0.06)	(0.01)	(0.01)
Exit Rate	12.51	0.62	10.84	0.12	12.74	0.67	10.85	0.21
First Stage F	*7.49*	*7.49*	*7.49*	*7.49*	*7.00*	*7.00*	*7.00*	*7.00*
Obs	*2,928*	*2,928*	*2,928*	*2,928*	*2,935*	*2,935*	*2,935*	*2,935*
FRL[30,70]								
Reduced Form	1.22	0.05	0.02	0.00	1.38 **	0.04	0.04 **	-0.01
	(0.81)	(0.07)	(0.02)	(0.01)	(0.70)	(0.07)	(0.02)	(0.01)
Exit Rate	13.06	0.62	10.85	0.12	13.21	0.67	10.86	0.21
First Stage F	*4.28*	*4.28*	*4.28*	*4.28*	*4.42*	*4.42*	*4.42*	*4.42*
Obs	*2,243*	*2,243*	*2,243*	*2,243*	*2,272*	*2,272*	*2,272*	*2,272*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification). All schools in participating districts with FRL data.

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at school-level are in parentheses.

Exit Index is the predicted probability of exit in the following year from a regression of school exit on experience, MA+ indicator, and log base salary.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

Table A1 - Math Achievement RD Results Excluding Districts with Possible TIF Treatment Contamination in 2012-2013

	2011-2012			2012-2013		
	All TIF-Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond	All TIF-Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage	0.39 ***	0.36 ***	0.35 **	0.44 ***	0.42 ***	0.43 ***
	(0.13)	(0.13)	(0.14)	(0.13)	(0.14)	(0.15)
Reduced Form (Elgibility)	-0.04	-0.04	-0.06	-0.09 *	-0.09 *	-0.09 *
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
First Stage F	*9.12*	*7.56*	*6.54*	*11.04*	*9.42*	*8.68*
Obs	*41,341*	*39,196*	*35,227*	*40,530*	*38,463*	*35,103*
Specifications						
Student Characteristics	X	X	X	X	X	X
Grade Fixed Effects	X	X	X	X	X	X
Test Score Mean 2009-10 (Grade x School)	X	X	X	X	X	X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Lebanon school district ceased its participation in Spring 2013.

In 2012-2013, Lebanon and Redmond school districts participated in the SB 252 Pilot program connected to Oregon's NCLB Waiver application

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table A2 - Reading Achievement RD Results Excluding Districts with Possible TIF Treatment Contamination in 2012-2013

	2011-2012			2012-2013		
	All TIF-Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond	All TIF-Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage	0.39 ***	0.37 ***	0.38 ***	0.41 ***	0.39 ***	0.40 ***
	(0.13)	(0.13)	(0.14)	(0.13)	(0.13)	(0.14)
Reduced Form (Elgibility)	0.10 **	0.11 **	0.10 *	0.05	0.05	0.06
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)
First Stage F	*9.36*	*7.96*	*7.19*	*9.89*	*8.50*	*7.95*
Obs	*41,199*	*39,297*	*35,473*	*40,466*	*38,650*	*35,297*
Specifications						
Student Characteristics	X	X	X	X	X	X
Grade Fixed Effects	X	X	X	X	X	X
Test Score Mean 2009-10 (Grade x School)	X	X	X	X	X	X

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Lebanon school district ceased its participation in Spring 2013.

In 2012-2013, Lebanon and Redmond school districts participated in the SB 252 Pilot program connected to Oregon's NCLB Waiver application

First Stage regressions: $TIF = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

Standard errors clustered at the school level are in parentheses.

Table A3 - RD Estimates of TIF Impact on School Exit Likelihood of Teachers - Excluding New/Closed Schools

	School Year							
	2010-11		2011-12		2012-13		2013-14	
	(1)	(2)	(2)	(3)	(3)	(4)	(4)	(5)
A) 2009-10 Full-Time Teachers								
Reduced Form	0.00	0.00	-0.06	-0.07	-0.07	-0.08 *	-0.10 *	-0.10 **
	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
First Stage	0.36 ***	0.35 ***	0.36 ***	0.35 ***	0.36 ***	0.35 ***	0.36 ***	0.35 ***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate	0.13	0.13	0.27	0.27	0.36	0.36	0.44	0.44
First Stage F	*8.55*	*8.61*	*8.55*	*8.61*	*8.55*	*8.61*	*8.55*	*8.61*
Obs	*3,147*	*3,147*	*3,147*	*3,147*	*3,147*	*3,147*	*3,147*	*3,147*
B) 2010-11 Full-Time Teachers								
Reduced Form			-0.07 *	-0.07 *	-0.10 **	-0.09 **	-0.13 ***	-0.13 ***
			(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
First Stage			0.34 ***	0.33 ***	0.34 ***	0.33 ***	0.34 ***	0.33 ***
			(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate			0.20	0.20	0.29	0.29	0.39	0.39
First Stage F			*7.68*	*7.57*	*7.68*	*7.57*	*7.68*	*7.57*
Obs			*3,164*	*3,164*	*3,164*	*3,164*	*3,164*	*3,164*
Specification								
Teacher Covariates		*X*		*X*		*X*		*X*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification).

Reduced Form regressions: $y = a + b*(z>=0) + c*z*(z>=0) + d*z*(z<0)$

First Stage regressions: $TIF = p + q*TIF + r*z*(z>=0) + s*z*(z<0)$

Data includes all base year (2009-10 or 2010-11) full-time (FTE>0.9) teachers in non-charter schools in participating districts with FRL data.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

School Exit indicator is equal to one in year t when a teacher that was observed working FT in school s in the base year (2009-10 or 2010-11) is no longer observed working FT in school s in year t.

Teacher covariates include: cubic in experience, log base salary, indicator for MA or above, indicator for Special Ed, indicator for Elementary School teacher.

Standard errors clustered at school-level are in parentheses.

Table A4 - RD Estimates of TIF Impact on Year-over-Year School Exit Likelihood of Teachers - Exlcuding New/Closed Schools

	School Year									
	2009-10		2010-11		2011-12		2012-13		2013-14	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reduced Form	0.06	0.06	0.00	0.00	-0.07 *	-0.07 *	-0.02	-0.02	-0.04	-0.06 *
	(0.04)	(0.04)	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	(0.02)	(0.03)	(0.03)
First Stage	0.37 ***	0.37 ***	0.36 ***	0.35 ***	0.34 ***	0.33 ***	0.33 ***	0.32 ***	0.34 ***	0.34 ***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Exit Rate	0.16	0.16	0.13	0.13	0.20	0.20	0.14	0.14	0.16	0.16
First Stage F	*8.77*	*8.78*	*8.55*	*8.61*	*7.68*	*7.57*	*7.81*	*7.67*	*8.03*	*7.97*
Obs	*128*	*128*	*127*	*127*	*128*	*128*	*128*	*128*	*128*	*128*
Specification										
Teacher Covariates		*X*		*X*		*X*		*X*		*X*

Each entry represents the coefficient estimate of interest from a different RD regression (linear specification).

Reduced Form regressions: y = a + b*(z>=0) + c*z*(z>=0) +d*z*(z<0)

First Stage regressions: TIF= p + q*TIF + r*z*(z>=0) +s*z*(z<0)

Data includes all base year (2009-10 or 2010-11) full-time (FTE>0.9) teachers in non-charter schools in participating districts with FRL data.

TIF-participating districts: Bend-Lapine, Crook, Greater Albany, Lebanon, Salem-Keizer.

School Exit indicator is equal to one in year t when a teacher that was observed working FT in school s in the prior year is no longer observed working FT in school s in year t.

Teacher covariates include: cubic in experience, log base salary, indicator for MA or above, indicator for Special Ed, indicator for Elementary School teacher.

Standard errors clustered at school-level are in parentheses.