Do Bonuses Affect Teacher Staffing and Student Achievement in High-Poverty Schools? Evidence from an Incentive for National Board Certified Teachers in Washington State

James Cowan and Dan Goldhaber

Center for Education Data & Research
University of Washington Bothell

March 20, 2015

The Center for Education Data & Research
University of Washington Bothell
3876 Bridge Way N. Ste. 201 Seattle, WA 98103
(206) 547-5585 cedr@u.washington.edu
www.cedr.us

Abstract

We study a teacher incentive policy in Washington State that awards a financial bonus to National Board Certified Teachers who teach in high-poverty schools. We study the effects of the policy on student achievement and teacher staffing using a regression discontinuity design that exploits the fact that eligibility for the bonus is based on the percentage of a school's population that is eligible for free- or reduced-price lunches. We find that the bonus policy increased the proportion of National Board Certified Teachers in bonus-eligible schools, through increases in both the number of existing NBCTs hired and the probability that teachers at these schools apply for certification. However, we do not find evidence that the bonus resulted in detectible effects on student test achievement.

Acknowledgments

This study was funded by the Bill and Melinda Gates Foundation and by the National Center for Analysis of Longitudinal Data in Education Research (CALDER), funded through grant #R305A060018 to the American Institutes for Research from the Institutes of Education Sciences, U.S. Department of Education. We thank both funders for their generous financial support. We additionally thank Nirav Mehta, Ben Ost, and participants at the 2015 AEFP Conference for helpful comments. Christopher Tien provided expert research assistance. We also thank the Washington Office of the Superintendent of Public Instruction (OSPI) and the National Board for Professional Teaching Standards (NBPTS) for providing the data used in this study. The views expressed are those of the authors and should not be attributed to the American Institutes for Research, its trustees, or any of the funders or supporting organizations mentioned herein. Any errors are attributable to the authors.

It is well-documented that teachers are one of the most important school-based influences on student learning. Yet, as with other schooling resources, there is substantial evidence that teaching quality is not equitably distributed across schools. A long literature on teacher mobility suggests that teachers are more likely to exit high-poverty schools (Borman and Dowling, 2008). Higher teacher turnover may result in a more inexperienced teaching staff and disrupt collaboration and school management (Clotfelter et al., 2011; Ronfeldt et al., 2013). These schools may also have greater difficulty attracting high-quality teachers for open positions (Jacob, 2007). As a consequence, high-poverty schools tend to have less qualified and less effective teachers (Chetty et al., 2014a; Goldhaber et al., 2014; Sass et al., 2012).

In response to these facts, policymakers have implemented a number of teacher recruitment and retention strategies for high-poverty schools. Unlike other common teacher incentive schemes, which target individual or group teacher effort, these school-based bonuses are intended to increase the supply of effective teachers to particular schools. We study a bonus policy for National Board Certified Teachers (NBCTs) in high-poverty schools in Washington State. Since the 2007-2008 school year, Washington has awarded a $5,000 bonus to NBCTs who teach in schools whose enrollment of free lunch eligible students exceeds a certain threshold. Given that NBCTs tend to be more effective at increasing student test scores than other teachers (Clotfelter et al., 2007, 2010; Cowan and Goldhaber, 2015; Goldhaber and Anthony, 2007), this policy specifically targets high quality teachers in low-SES schools.

We study the effects of the NBCT bonus policy on student achievement and teacher staffing outcomes using a regression discontinuity design. We provide some of the first evidence of the student achievement effects of teacher retention policies. We find that the bonus policy increased the number of NBCTs in high-poverty schools, through increases in both the number of existing NBCTs hired and the probability that teachers at bonus-eligible schools applied for certification. There is suggestive evidence of an improvement in retention of NBCTs; however, statistical significance is not robust to variations in model specifications. Despite the evident improvements in teacher staffing, we do not find positive student achievement effects from the bonus policy. Our estimated treatment effects are mostly negative, albeit imprecise and statistically insignificant.

The combination of the positive teacher staffing results and null student achievement effects may be surprising. However, the student achievement effects of a teacher incentive policy will depend both on the effectiveness of targeted teachers and changes in teacher labor supply. Given the responsiveness of NBCTs to the bonus and the magnitude of the differences in teacher effectiveness between NBCTs and non-NBCTs commonly estimated in the literature, it does not appear that the student achievement and teacher staffing effects we estimate are incongruous.

I Background and Literature Review

The Washington bonus program, called the Challenging Schools Bonus (CSB), is one of many attempts to use financial incentives to improve the quality of teaching in high-needs schools. Initiatives like this are designed not only to address the fact that students in high-poverty schools often need additional support to succeed academically, but also because a number of studies show that teacher quality, whether measured by observable credentials or value-added to test scores, is inequitably distributed across schools (Clotfelter et al., 2011; Goldhaber et al., 2014; Sass et al., 2012). Boyd et al. (2013) suggest that these empirical findings at least partially represent the preferences of teachers for certain kinds of working environments. In other industries where salaries are less centrally determined, firms with undesirable job characteristics are forced to provide compensating differentials to attract qualified workers. Targeted, or "market-based," incentive policies, therefore seek to address teacher preferences for school characteristics within the confines of a codified teacher salary schedule.

The program we study rewards both obtaining a particular teaching credential and service in high-needs schools. The National Board for Professional Teaching Standards (NBPTS) is a national organization that offers voluntarily certification for accomplished teaching. Nationally, more than 110,000 teachers have earned Board certification (National Board for Professional Teaching Standards, 2014b). NBPTS is broadly seen as a credential identifying teacher excellence, and prior work has found that NBCTs are more effective at raising student test scores than other teachers (Clotfelter et al., 2007, 2010; Cowan and Goldhaber, 2015; Goldhaber and Anthony, 2007).[1] In addition to the general bonus NBCTs in Washington receive, the CSB program allocates an additional bonus to NBCTs serving in high-need schools, defined based on the free- and reduced-price lunch eligibility level of the school's students. As with other indicators of teacher quality, NBCTs appear to be inequitably distributed across schools (Goldhaber et al., 2007; Humphrey et al., 2005). These targeted financial incentives represent one strategy for addressing observed teacher quality gaps. By increasing the desirability of teaching in high-needs schools, they may both improve teacher retention and the quality of the pool of applicants for open teaching positions. The effectiveness of targeted incentive policies therefore relies on both the sensitivity of incumbent teachers' mobility decisions to increases in salary and the relationship

[1] However, not all studies have found statistically significant, positive effects (e.g., Harris and Sass, 2009).

between the potential for additional compensation and teachers' job application decisions.

The literature on teacher mobility suggests that turnover may be particularly acute in high-needs schools. A long literature on teacher turnover suggests that teachers are more likely to leave schools with larger disadvantaged populations (Borman and Dowling, 2008). Targeted financial incentives may therefore increase the desirability of working in high-needs schools relative to other employment options. Several studies have found that school districts with higher salaries have lower teacher mobility (Hanushek et al., 2004; Imazeki, 2005; Lankford et al., 2002). However, Hanushek et al. (2004) and Imazeki (2005) conclude that relatively large salary premia would be required to equalize attrition rates between urban and suburban schools.

While the literature suggests that teacher exits are sensitive to salary differentials, there is less consistent evidence on the role of teacher pay in attracting desirable teacher candidates. For a variety of reasons, teachers may be less responsive to recruitment incentives than retention incentives of the same dollar amount. Job changes necessarily involve transaction costs in the form of interview and moving expenses and lost school-specific human capital.[2] Studies examining variation in teacher salaries within states over time have generally found positive effects on long-term student outcomes (Card and Krueger, 1992; Loeb and Page, 2000). Student achievement and high school graduation effects may be due to improved teacher hiring, retention, or other factors and evidence from analyses of local labor markets that focuses on teacher credentials is mixed. Hanushek et al. (2005) examine the mobility patterns of a small sample of teachers by their observable credentials and measured value-added and find little evidence that districts with higher salaries attract the most effective mobile teachers. Hanushek et al. (1999) find that salary increases are associated with improvements in the teacher licensure test scores of new candidates, although this only holds in larger districts. They also find positive effects on student achievement and salary increases, although these effects do not appear to be explained by changes in teacher experience or turnover. Finally, Figlio (2002) finds that non-unionized school districts that raise their salaries relative to other nearby districts increase their hiring of teachers from selective colleges and with subject-specific majors. However, this does not appear to hold for unionized school districts.

While high-quality teachers are not equally represented in low-achieving and high-poverty schools, this does not appear to be driven by higher rates of attrition among high-performing teachers. The available evidence suggests that high value-added teachers are less likely to depart their schools generally (Boyd et al., 2008; Goldhaber et al., 2011;

[2]For instance, Glazerman et al. (2013), which is discussed in more detail below, find that only 22% of teachers recruited for a $20,000 recruitment incentive even applied for the program.

Hanushek et al., 2005). Boyd et al. (2008) find that this relationship also holds for effective teachers in low-performing schools. Instead, sorting among mobile teachers may contribute to differences in teacher effectiveness across schools. Boyd et al. (2008) find that more effective teachers who do leave are more likely to depart for higher-achieving schools. Similarly, controlling for a variety of teacher and school characteristics, Clotfelter et al. (2011) find that NBCTs are about 13% less likely to leave their school than other teachers. While NBCTs may be less likely to depart than other teachers, possession of the certificate may increase mobility by providing teachers a verifiable signal of teacher quality. In a study assessing the mobility patterns of participants in the NBPTS certification process, Goldhaber and Hansen (2009) find that NBCTs are significantly more likely to depart their schools than unsuccessful applicants.

While there is a relatively long literature linking teacher turnover to working conditions and salary, there is less evidence about the effects of particular policy interventions. Only a few studies have examined the effects of targeted compensation policies on teacher staffing.[3] The policies that have been studied fall under two general types. First are recruitment bonuses, which offer teachers a one-time financial incentive to take positions in a certain class of schools. The second type of policy is a recurring salary increase, which offers generally smaller but continuous incentives so long as teachers work in an eligible school. The policy we study is of the latter type.

Two studies of recruitment bonuses have found that they increase the likelihood that teachers take positions in eligible schools but not that they improve retention. Glazerman et al. (2013) analyze a policy that provided a randomized group of schools the opportunity to offer high-performing teachers $20,000 bonuses to transfer to a low-achieving school for at least two years. They find clear evidence of recruitment effects: teachers recruited to eligible positions were 43 percentage points less likely to have less than six years of experience and more than twice as likely to possess National Board certification. While they find that turnover is lower for recruited teachers during the two-year window, they find no effects on retention thereafter. Steele et al. (2010) study a policy initiative in California, the Governor's Teaching Fellowship, which conditioned a $20,000 scholarship on teaching in a low-performing school for four years. As with the findings from the Teacher Transfer Initiative, they find a substantial increase in the likelihood that targeted teachers work in such schools, but that attrition rates are similar to other teachers in the same schools.

Continuing bonus policies, which provide an ongoing incentive to work in high-needs

[3]Here we differentiate recruitment and retention bonuses, which link financial incentives to serving in a particular group of schools, from general salary increases or performance bonuses, which link salary increases to increased effort or effectiveness. See Podgursky and Springer (2007) for a review of the latter and Dee and Wyckoff (2013) for a recent study examining the link between performance pay and teacher turnover.

schools, may be expected to improve both recruitment and retention of teachers. While we are not aware of studies that have examined the recruitment effects of continuing, targeted incentive policies, a few studies have found that they increase teacher retention. Clotfelter et al. (2008) study a temporary North Carolina program that awarded $1,800 bonuses to math, science, and special education teachers in high-poverty or low-achieving schools. They find that the bonus policy reduced the turnover of targeted teachers by about 17%. Springer et al. (2014) assess a Tennessee program that paid $5,000 bonuses to highly-rated teachers in low-achieving schools and find that receipt of the bonus improved teacher retention.[4]

In sum, prior research suggests that financial incentives may improve both teacher recruitment and retention in targeted schools. It does not appear that one-time awards that incentivize teachers to work in particular schools improve teacher retention and the total long-term effects of such policies are therefore unclear. On the other hand, retention bonuses may improve both teacher turnover and teacher recruitment at eligible schools. As the retention bonus literature has so far focused on the former effect, the existing evidence may understate the total effect on school staffing.

Research on teacher staffing and National Board certification suggests a few channels through which the Challenging Schools Bonus could affect student achievement. First, high-poverty schools appear to have lower average teacher effectiveness than lower-poverty schools (Goldhaber et al., 2014; Sass et al., 2012). If bonus policies help such schools retain or attract effective teachers, they may, over time, increase the average effectiveness of teachers at high-needs schools. Although the empirical research on the effectiveness of NBCTs relative to other teachers is not uniformly positive, most of the available literature suggests that a bonus policy using this signal of teacher quality should identify effective teachers (Clotfelter et al., 2007, 2010; Cowan and Goldhaber, 2015; Goldhaber and Anthony, 2007; Harris and Sass, 2009). Second, teacher turnover may itself be harmful to student achievement, and policies that improve retention may have spillover effects on other students in the same school (Ronfeldt et al., 2013). Third, past research has documented positive peer effects from high value-added or collaborative teachers (Jackson and Bruegmann, 2009; Raudenbush et al., 1992). To the extent that NBCTs are more effective or more collaborative peers, they may improve the productivity of other teachers in their schools. Although there are no studies we are aware of documenting peer spillover effects for NBCTs specifically, Loeb et al. (2006) survey NBCTs in Washington and find that more

[4]On the other hand, Fulbeck (2014) analyzes teacher mobility in Denver under the Professional Compensation System for Teachers, which includes a bonus for teachers who work in schools with high poverty rates, and finds that teachers in eligible schools are more likely to leave their school than other teachers receiving bonuses.

than 50% reported increasing their school or district leadership activities following certification. Finally, the policy also incentivizes incumbent teachers to participate in the NBPTS assessment process in order to become eligible for the bonus. The effect of participation on teacher effectiveness is not clear. On the one hand, the assessment process requires 200-400 hours of time to complete (Boyd and Reese, 2006). Applicants may therefore have less time to devote to classroom activities and be less effective at classroom teaching during the submission year (Goldhaber and Anthony, 2007).[5] On the other hand, the NBPTS application process includes a number of assessment and reflection activities that may improve teaching practice.[6] More generally, while there are clear theoretical reasons to believe that bonus policies may increase student achievement, there is little empirical evidence on this question.

Washington's Challenging Schools Bonus Program

Washington State has awarded a salary incentive for National Board Certified Teachers since the 1999-2000 school year. Initially set at 15% of salary, the state fixed the bonus at $3,500 per year in 2000 and raised it to $5,000 in 2007. At the same time, Washington introduced an additional bonus for teachers in high-poverty schools. The new program, called the Challenging Schools Bonus (CSB), awards a bonus of up to $5,000 to National Board teachers at high poverty schools in addition to the bonus all NBCTs in Washington State receive.[7] The state covers the cost of the bonus, which is paid as a lump sum to teachers during the summer following the school year.

The CSB is awarded to teachers who possess the NBPTS teaching credential and work in schools with sufficiently high enrollments of students eligible for free- or reduced-price lunch (FRL) programs. The definition of high-poverty varies across years and school level. The particulars of the policy are described in **Table 1**. During the first year of the program, schools with FRL enrollment exceeding 70% of total enrollment were eligible for the bonus. Thereafter, eligible schools were elementary schools with FRL enrollment exceeding 70% of total enrollment, middle schools with FRL enrollment exceeding 60%, and

[5]However, Harris and Sass (2009), who estimate models with a full range of teacher, school, and student fixed effects using Florida data, find statistically significant, negative effects on current participation only on one of the four tests they consider.

[6]Sato et al. (2008) find that participation improves teachers' assessment skills. However, research that has studied this question by examining within-teacher variation in student achievement has generally not found that teachers are more effective following certification (Chingos and Peterson, 2011; Goldhaber and Anthony, 2007; Harris and Sass, 2009).

[7]The amount of the bonus is prorated by the proportion of time teachers spend in an eligible school. Starting with the 2011-2012 school year, the amount of the bonus was reduced to $3,000 for the first year a teacher earns National Board certification.

high schools with FRL enrollment exceeding 50%. Besides the change after the first year of the policy from a constant threshold of 70% across all school levels to one that varied by level, two features of the eligibility criterion vary across years. In the earliest years of the bonus, OSPI calculated eligibility based on the maximum of the FRL enrollment reported in the state's Core Student Record System (CSRS) and the Childhood Nutrition Services (CNS) report.[8] For the 2011-2012 school year and beyond, OSPI used only the October count from the state enrollment system. Second, between 2008 and 2010, OSPI grandfathered schools into the Challenging Schools Bonus based on their eligibility in the previous year. This practice was discontinued for the 2011-2012 school year.

Two prior studies have described patterns of mobility under the CSB program. Using similar data sources, Elfers and Plecki (2014) and Simpkins (2011) document that the proportion of NBCTs in high-poverty schools increased following the adoption of the CSB in 2007-2008. By 2013, CSB-eligible schools accounted for 28% of all teaching positions and 31% of all NBCTs. Both studies also find an increase in the likelihood that teachers in high-poverty schools apply for Board certification following the introduction of the bonus. The two studies reach somewhat divergent conclusions about changes in recruitment and retention under the CSB program. Elfers and Plecki (2014) find that retention rates for NBCTs in high-poverty schools are similar to those in low-poverty schools and slightly higher than other teachers in high-poverty schools. On the other hand, Simpkins (2011) finds that the number of teachers departing Challenging Schools exceeds the number of teachers entering.

While the descriptive evidence points to an increase in NBCTs in high-poverty schools following the introduction of the policy, the causal contribution of the program is not clear. As documented above, there is empirical evidence that retention rates of NBCTs and of teachers at high-poverty schools may differ from other teachers with similar characteristics. The CSB policy may therefore increase the number of NBCTs in targeted schools by reducing attrition even if the net migration is negative. The introduction of the incentive policies also coincided with a major effort to increase the number of NBCTs in Washington State through conditional loans for application fees and candidate support and networking initiatives (Elfers and Plecki, 2014; National Board for Professional Teaching Standards, 2010). The number teachers earning certification each year increased from 49 in 2000 to 1,251 in 2008 (Plecki et al., 2010). By 2014, Washington was producing by far the most new NBCTs nationwide (National Board for Professional Teaching Standards, 2014a). Consequently, the number of NBCTs in Washington State increased substantially

[8] The Comprehensive Education Data and Research System (CEDARS) replaced CSRS in 2010. We refer to this as the "CSRS" measure throughout for simplicity.

in both CSB-eligible and CSB-ineligible schools during the same time period as the policy was introduced. This pattern is reflected in **Figure 1**, which plots the proportion of teachers possessing the NBPTS credential by 2006-2007 school FRL enrollment in 2007-2008 (the first year of the bonus policy) and 2012-2013. While the share of NBCTs increased more rapidly in high-poverty schools, there is a substantial increase in ineligible schools as well. Disentangling the effects of the bonus program from unobserved school characteristics and statewide trends is the key methodological challenge in assessing the CSB program.

II Data and Empirical Methods

Empirical Methods

The Challenging Schools Bonus seeks to improve the staffing of schools that predominantly serve low-income students by improving their staffing of effective teachers. In this study, we assess the extent to which the bonus improves school staffing and increases average student achievement. The key difficulty in assessing the effects of a policy like the CSB is that assignment is correlated with a school characteristic, poverty, that is known to be associated with both student achievement and teacher recruitment and retention.[9] Therefore, comparisons of student achievement and staffing outcomes at eligible and ineligible schools may be biased by unobserved factors.

To identify the effect of the bonus policy on student achievement and teacher staffing, we exploit the fact that eligibility is based on a discontinuous function of school-wide FRL enrollment. The discontinuous eligibility rule allows us to employ a regression discontinuity design (RDD) as our primary research strategy. The RDD relies on this sharp break in the relationship between the FRL measure (the "forcing variable") and eligibility status to estimate the treatment effect as the difference in average outcomes between schools just eligible for the CSB policy and just ineligible for the CSB policy.[10] By focusing on changes in outcomes at the eligibility threshold, the regression discontinuity approach ignores variation in outcomes that may be associated with factors correlated with school poverty but are not caused by the program itself. The cost of this approach is that our estimated treatment effects only describe outcomes for schools near the eligibility threshold.[11] In particular, our

[9] See Borman and Dowling (2008) for a meta-analysis assessing the literature on school characteristics and teacher mobility. Ehlert et al. (2013) consider the role of poverty in assessing school effectiveness.

[10] In practice, we employ local linear regressions, which model the relationship between FRL and school outcomes in the neighborhood of the eligibility threshold, as naive estimates that rely solely on comparisons of average outcomes perform less well in finite samples (Hahn et al., 2001).

[11] In the evaluation literature, these are known as local average treatment effects, as they describe treatment effects for a certain subset of the population. In our primary specification, the local average treatment effects describe treatment effects for schools at the eligibility threshold. As we discuss in greater detail below,

results may not generalize to schools with greater shares of FRL enrollment.

Due to the grandfathering provision, during the early years of the bonus program, eligibility in *future years* also varies discontinuously at the eligibility threshold. Therefore, there is a discontinuity in teachers' expectations of both current and future salary at the eligibility threshold. Because of the grandfathering provision, we can construct a measure of current and past FRL enrollment that generates both a sharp discontinuity in contemporaneous treatment as well as a discontinuity in the length of exposure to the CSB program. Given the eligibility rule, we construct a measure of current and past FRL enrollment that completely determines CSB eligibility.[12] Following the criteria in Table 1, we generate the forcing variable

$$x_{st} = \begin{cases} \max(FRL_{st}^{CSRS}, FRL_{st}^{CNS}) - c_{st} & \text{if } t = 2008 \\ \max(FRL_{st}^{CSRS} - c_{st}, FRL_{st}^{CNS} - c_{st}, x_{st-1}) & \text{if } t = 2009 \\ \max(FRL_{st}^{CSRS} - c_{st}, x_{st-1}) & \text{if } 2010 \leq t \leq 2011 \\ FRL_{st}^{CSRS} - c_{st} & \text{if } t \geq 2012 \end{cases} \quad (1)$$

where FRL_{st}^{CSRS} is the Core Student Record System FRL enrollment share, FRL_{st}^{CNS} is the Child Nutrition Services FRL enrollment share, x_{st-1} is the prior year's value of the forcing variable, and c_{st} is a school-specific threshold that depends on the school year and the grades offered in the school. As we discuss below, we uncover some evidence of manipulation of the CNS measure of school poverty during the year that the grandfathering provision is introduced. We therefore additionally generate a forcing variable that uses only the information about school eligibility contained in the CSRS measure, which we show to be less susceptible to noncompliance. Using the forcing variable x_{st}, we estimate

$$Y_{ist} = \delta CSB_{st} + f(x_{st}) + Z_{ist}\beta + \epsilon_{ist} \quad (2)$$

for 2008-2013, the years in which the CSB program is in effect. We estimate $f()$ by local linear regression with a triangular kernel. Hence, in practice, we estimate Eq. (2) by weighted least squares for schools with x_{st} within 0.35 of the eligibility threshold c_{st},

we additionally estimate models with an alternative version of the forcing variable that uses only the CSRS enrollment. In this case, the local average treatment effect describes the effect of the CSB program for schools at the eligibility threshold for which eligibility is decided by the CSRS measure alone.

[12]In the data provided by OSPI, there are three schools whose reported CSB status in one year did not correspond either to the reported FRL enrollment data or to eligibility reports from other years. Because we could not verify whether these schools actually received the bonus, we dropped them from our study. The findings reported below, however, are not sensitive to their inclusion in the sample.

which is the median of the optimal bandwidths among the outcomes we consider.[13] We estimate specifications with a number of additional covariates in Z_{ist}. In all models, these include school grade level, school racial composition, current FRL enrollment share, log enrollment, 2007 mean student achievement in math and reading, and mean 2007 teacher NBCT status, experience, advanced degree possession, and turnover.[14] In the student-level models, we additionally control for student gender, race, learning disability status, participation in gifted, bilingual, special education, or Learning Assistance Program services, FRL eligibility, and 2007 mean test scores in the given grade and school.[15] For students who were enrolled in grades 3-8 during the 2006-2007 school year, we have pre-treatment student achievement measures. As including measures of lagged achievement may improve the precision of estimates, we additionally estimate the regression discontinuity models on this sample of students controlling for 2007 math and reading achievement (interacted with grade and year).

It is possible that the effects of the CSB policy could accumulate over time. This may be particularly true for the student achievement outcomes, where attendance at a more effective school may be expected to increase student achievement by a given amount for each year of attendance. The regression discontinuity estimates therefore represent a weighted average of the effects of differing lengths of exposure. Given this, we estimate a specification where we treat the eligibility rule as an instrument for the number of years of CSB eligibility. Specifically, we estimate

$$\begin{aligned} Y_{ist} &= \delta CSB_{st} + f(x_{st}) + Z_{ist}\beta + \epsilon_{ist} \\ CSB_{st} &= \gamma 1(x_{st} \geq 0) + g(x_{st}) + Z_{ist}\pi + \eta_{ist} \end{aligned} \tag{3}$$

by two stage least squares (2SLS). The coefficient on CSB_{st} in Eq. (3) divides the discontinuity in outcomes at the eligibility threshold by the discontinuity in the duration of treatment. For the student achievement regressions, we indicate the number of years the student has attended a CSB-eligible school. For the teacher staffing measures, we indicate the number of years the current school has been eligible for the bonus policy.

The key identifying assumption in the regression discontinuity design is that unob-

[13]We estimate optimal bandwidths using the modified cross-validation algorithm described by Lee and Lemieux (2010). Imbens and Kalyanaraman (2012) suggest an analytic optimal bandwidth for the regression discontinuity design. We prefer the cross-validation approach because our design relies on a treatment and forcing variable that are fixed at the school level and the Imbens-Kalyanaraman bandwidth calculation assumes observations are independent and identically distributed. We return to the sensitivity of the RD estimates to bandwidth choice below.

[14]Recall that CSB eligibility relies on *lagged* FRL enrollment share.

[15]The Learning Assistance Program is a state-financed program of supplemental learning support for students deficient in reading, writing, and math.

served factors affecting teacher staffing and student achievement vary continuously around the eligibility threshold. If this assumption holds, then the only reason we should observe a sharp break in school outcomes is due to the change in eligibility status for schools just above the FRL cut-offs. This assumption may fail if other characteristics of schools vary discontinuously at the eligibility threshold, such as if other programs condition eligibility based on similar rules.[16] It may also fail if schools have some discretion over their enrollment counts such that they can manipulate their eligibility status. Below, we document some evidence of possible manipulation in the enrollment variables. This suggests that simple regression discontinuity models may provide biased estimates of treatment effects. However, as we observe the introduction of the policy, we can directly control for pre-treatment measures of the outcome variables. This allows us to show that estimates are robust to assuming only that *changes* in school outcomes, rather than levels of school outcomes, are uncorrelated with treatment assignment conditional on the forcing variable. We also show that our estimates are robust to alterative definitions of the forcing variable.

Data

In this study, we use data on student assessments and teacher staffing from two databases maintained by the Office of the Superintendent of Public Instruction (OSPI). We construct data on teacher turnover and credentials using the S-275, which is the employment reporting system for public schools in Washington State. Districts report school employees who have an employment contract in place by October 1 of each school year. Beyond school of employment, the S-275 additionally includes information about employees' assignments and teaching credentials. These include educational attainment, experience, salary, and teacher demographics. Using this database, we construct a panel of teachers for the schools years 2007-2008 through 2012-2013. The S-275 reports assignment information that includes the total full-time equivalency (FTE) in each reported position. We define teachers as those with base contracts that have positive FTE in classroom teaching positions and assign teachers to a school if their reported time in that building exceeds 50% of their total FTE.[17] Using this assignment rule, we include all teachers working in Washington public schools that qualify to be considered for CSB eligibility.[18] The remaining sample repre-

[16] We are not aware of any such programs.

[17] About 18% of all teaching staff reports a position in more than one building. However, this includes teachers with an assignment reported at the district office and may overstate the number of teachers who work in more than one school.

[18] We omit schools that ever enroll fewer than 30 students, which are generally not eligible to to qualify for the CSB, as well as schools with no K-12 enrollment in any year or schools missing enrollment information required for determination of eligibility. The final sample includes 1,718 schools.

sents about 90% of public school employees who are reported as working in classroom teaching positions. As we are primarily interested in how the bonus policy affects teacher retention, we define a teacher as moving if she is not assigned to the same school in the S-275 during the following academic year.[19] This includes teachers who switch schools and those who exit the public school system.

We supplement the S-275 with data on National Board candidates provided by NBPTS. This data includes the year of submission and assessment results of all applications originating from public or private schools in Washington State. Our measure of National Board certification therefore only includes teachers who initially obtained certification in Washington and may understate the total number of NBCTs. To the extent that the bonus policy incentivizes cross-state migration to eligible schools, our estimates of staffing effects would likely be downward biased. However, cross-state teacher mobility appears to be relatively rare. During the 2011-2012 school year, while nearly 7% of teachers had worked in a different public school in the same state during the previous year, only 0.5% of teachers had worked in a public school in another state (National Center for Education Statistics, 2015). There are additionally very few NBCTs in Oregon and Idaho, the two states that share a border with Washington (National Board for Professional Teaching Standards, 2014b). Given the relative rarity of both cross-state mobility and National Board certification, the degree of measurement error in our NBCT indicator is likely to be small.

We present summary statistics for teachers across all years of our sample in **Table 2**. In CSB schools, 14% of teachers depart each year, which is fairly similar to the statewide average of 13%. However, we do find some differences in observable teacher credentials: CSB teachers have about an average of one year less of teaching experience and are about 4 percentage points less likely to have an advanced degree. As a first indication that the CSB program may influence the supply of teachers, we find that 10% of CSB teachers in our sample possess an NBPTS teaching certificate, compared to the sample average of 8%.

The student-level data includes student demographic information, annual math and reading assessment data in grades 3-8 and 10, and information on participation in various special programs. Because our identification strategy ought to provide unbiased estimates of treatment effects without relying on pre-test information, we use all years for which testing outcomes are available. For 2008-2013, we use standardized reading assessments in grades 3-8 and 10 and standardized math assessments in grades 3-8. In 2010, the state replaced the 10th grade standardized math assessment with end-of-course assessments in

[19]The S-275 includes all teachers employed by October 1 of the reporting year, so our sample may exclude a small number of teachers who are laid off and subsequently recalled or who sign a contract after October 1 and stay at the school for only one year.

algebra and geometry. Given that the decisions about the timing of end-of-course assessments may be endogenous to bonus eligibility, we only consider outcomes on the standardized 10th grade math assessment from 2008-2010. We standardize all assessments by grade and year. Our main analytical sample includes 2,470,049 student-year observations in math and 2,711,038 in reading.

We conduct additional analyses using two subsamples of students. In some specifications, we use student-level assessments from before the implementation of the CSB as an additional control variable. In these cases, we limit the sample of students to those enrolled in tested grades in public schools during the 2006-2007 school year. We also use a sample of students matched to their math and reading teachers in grades 4-8. For 2008-2009, the classroom teacher is identified using the proctor listed on the end-of-year assessment. Because this teacher may be a homeroom or other teacher for middle school students, we limit this sample to elementary students and take other measures to ensure this sample only includes likely matches.[20] Because students in upper grades are more likely to have a homeroom or other teacher as a proctor, we only include students in grades 6-8 for years 2010-2013 in this sample.

We present summary statistics for students in **Table 3**. Unsurprisingly, CSB-eligible schools are quite different than the state as a whole. Student achievement is nearly 0.4 standard deviations lower in CSB schools than in other schools during the same time period. Students are also much more likely to be members of an underrepresented minority: 42% and 8% of students in CSB schools are Hispanic and African-American, respectively, compared to just 17% and 5% overall. Not surprisingly given these demographics, CSB students are nearly three times more likely to participate in bilingual programs than the state average.

Evidence of Noncompliance and the Validity of the Research Design

Before we proceed to the main results, we conduct two preliminary tests of the validity of the research design.[21] We first follow the approach of McCrary (2008b) and test for

[20] We only include matches where the listed proctor is a certified teacher in the student's school, has reported 1.0 full-time equivalency in that school, and is endorsed to teach elementary education. The data related to students and teachers used in this study are linked using the statewide assessment's "teacher of record assignment", a.k.a. assessment proctor, for each student to derive the student's "teacher". The assessment proctor is not intended to and does not necessarily identify the sole teacher or the teacher of all subject areas for a student. The "proctor name" might be another classroom teacher, teacher specialist, or administrator. For 2010, teachers are reported both as the exam proctor and with student and teacher scheduling administrative records. For this year, our proctor match agrees with the reported classroom teacher for about 95% of students in both math and reading.

[21] The regression discontinuity approach relies on the eligibility threshold generating random variation in receipt of the bonus. As an initial verification of the assumptions underlying the RDD, we assess the

discontinuities in the density of the forcing variable near the eligibility threshold. We then test for discontinuities in baseline outcomes and other student and school characteristics at the CSB eligibility threshold. Both tests suggest possible evidence of manipulation in the FRL enrollment measures that determine school eligibility.

If schools are able to influence their reported FRL-eligible enrollment by increasing efforts to recruit students or by misreporting enrollment, we may find that schools are unexpectedly more likely to fall just to the right of the eligibility threshold than just to the left. There is some evidence that this occurs with Title I funding, which also uses an eligibility rule that relies on FRL enrollment shares (Cascio et al., 2010; Matsudaira et al., 2012). Manipulation of the forcing variable potentially signals a violation of the key identifying assumption that unobserved differences in student or teacher staffing outcomes do not vary discontinuously at the eligibility threshold. In order to test for manipulation of the eligibility rule, we estimate the density of the forcing variable using the method of McCrary (2008b). In **Figure 2a**, we plot the estimated density of the forcing variable using the school-level data.[22] We estimate that the density of the forcing variable is about 23.0% (s.e. = 0.100) greater than expected to the right of the eligibility threshold and this difference is statistically significant. While significant, the degree of manipulation appears to be less than is the case with Title I, where Matsudaira et al. (2012) find a discontinuity in the distribution of FRL for a Northeastern school district of nearly 90%.

The forcing variable we use in this study is a composite of two FRL measures. To better understand the source of the discontinuity in the density, we estimate the density of each annual measure separately. In Figure 2b, we plot the kernel density estimate of the annual CSRS poverty measure for 2008-2013.[23] The estimated discontinuity at the eligibility threshold is 3.1% (s.e. = 0.078) and is not statistically significant. Next, we plot the annual CNS poverty measure for 2008-2009, the two years in which it determines CSB eligibility, in Figure 2c. While not statistically significant, the estimated discontinuity at the eligibility threshold, 18.3% (s.e. = 0.151), is estimated using fewer years of data and is nearly six times larger than with the CSRS measure. Finally, we construct a forcing variable equivalent of Eq. (1) that uses only the information about enrollment embedded in the CSRS poverty measure. The density of this measure is plotted in Figure 2d. Given the grandfathering provisions, the effect of the small discontinuity illustrated in Figure 2b

discontinuities in treatment at the eligibility thresholds. Using school-level data, we estimate that schools at the eligibility threshold have received the CSB for an average of 1.4 years longer than schools just below the threshold.

[22]We implement the density test using the Stata package DCdensity (McCrary, 2008a). Given the grouped nature of the data, we estimate standard errors using a clustered bootstrap on schools with 400 iterations.

[23]We define this as the prior year's FRL enrollment share centered around the school-year-specific eligibility threshold.

accumulates over time; however, the discontinuity, 10.0% (s.e. = 0.088), is less than half the size of that using the actual eligibility rule and is not statistically significant.

Analyzing discontinuities in baseline outcomes and student characteristics at the eligibility threshold, we find some evidence consistent with the density test of violations in the identifying assumptions. In Panel A of **Table 4**, we present the results of models that fit the regression discontinuity model to our data with 2007 school averages of the outcome variables as dependent variables. Using the primary forcing variable, we find that CSB-eligible schools had an average of 0.007 fewer NBCTs in 2007, which is statistically significant. When we instead use the alternative forcing variable that uses the CSRS poverty measure, the difference is -0.005 and no longer statistically significant. However, we do find a discontinuity at the CSRS threshold of -0.023 for advanced degree possession. None of the other outcomes is statistically significant. We next test for discontinuities in student covariates at the eligibility threshold in Panel B. Using the primary forcing variable, we find that students in eligible schools are about 0.5 percentage points more likely to utilize special education services (significant at the 10% level) and 2 percentage points more likely to be African American (significant at the 1% level). The latter result persists when we use the CSRS forcing variable instead. However, we do not find statistically significant differences in either school-wide or individual 2007 test scores at either eligibility threshold.

Given the evidence for possible noncompliance with the treatment rule, we conduct two primary robustness checks in the results that follow. First, we show that our regression discontinuity results are not sensitive to using the alternative definition of the forcing variable that exhibits less evidence of manipulation. Second, we additionally estimate all models with controls for school and student covariates. Because we observe the introduction of the policy, our control vector includes school-wide averages of all student and teacher outcomes from the 2006-2007 school year. This allows us to control for any pre-existing differences in staffing or student outcomes near the eligibility threshold. Therefore, we need only assume that changes in student and teacher outcomes are uncorrelated with treatment status once we control for the observed forcing variable.

III Effects of the Challenging Schools Bonus on School Staffing

The intention of the Challenging Schools Bonus is to improve the quality of instruction in high-needs schools. We therefore begin by characterizing the CSB treatment by its effects on school staffing. We graph local linear functions of the staffing outcomes against the forcing variable in **Figure 3** and present the regression results in **Table 5**. In Panel A, we estimate sharp regression discontinuity models using the primary eligibility rule and find that eligibility for the bonus increased both the total proportion of NBCTs and the proportion of teachers who are Board candidates. We estimate that the CSB increased the proportion of NBCTs by about 0.022 and the proportion of new candidates by about 0.007, which are both statistically significant at the 1% level. When we instrument the number of years with the eligibility indicator, we find an increase of about 0.015 per year of eligibility. For comparison, these effects represent about 25% of the baseline proportion of NBCTs and 50% of the annual baseline proportion of teachers submitting an application. The mean staff is about 30 teachers, which suggests an increase of about one half of an NBCT in eligible schools per year of eligibility. Importantly, these results are robust to the inclusion of school characteristics and pre-treatment outcomes in columns (2) and (4) and the use of the alternative forcing variable in Panel C. Reflecting the slightly lower levels of staffing by NBCTs in 2007 among eligible schools, the alternative estimates are slightly larger and suggest the CSB program increased the proportion of NBCTs by 0.026-0.030. The estimates for NBPTS applications are similar. Finally, we consider the effects on other teacher characteristics in columns (5) - (10). We find little evidence that the policy influenced the experience, educational attainment, or turnover of teachers in eligible schools.[24]

Although the National Board for Professional Teaching Standards offers certification in several areas, the most common certificates in Washington are for teachers in tested subjects and grade levels (Cowan and Goldhaber, 2015). To provide additional context for the student achievement effects, we therefore repeat our analysis with a sample of students in math and reading in grades 4-8 who are matched to their math or reading teacher. We present the results of this analysis in Panel D for math classrooms and Panel E for reading classrooms. The policy appears to have had greater effects on teacher staffing in tested classrooms. We estimate that eligibility increased the likelihood that students in math classes were assigned to an NBCT by 5-6 percentage points, while students in reading classrooms were about 3.5 percentage points more likely to have an NBCT. Eligibility appears to have only increased the application rate of teachers assigned to mathematics classrooms. We estimate that students in eligible schools are about 1.5 percentage points more likely to have an NBCT candidate as a teacher, which is statistically significant with the inclusion of covariates. Estimates for the reading classrooms are negative and not sta-

[24]We consider the sensitivity of these results to bandwidth choice in Appendix A. We plot estimated coefficients and 95% confidence intervals against bandwidth choice in Figure A.1. For both the NBPTS staffing outcomes, estimated effects are statistically significant for bandwidths above about 0.25. Using the optimal bandwidths for each outcome, we show in Table A.1 that we find statistically insignificant effects on overall NBCT staffing, although estimates are statistically significant when covariates are included. Results for NBPTS applications are similar to those presented here.

tistically significant. As with the estimates using the main teacher sample, we find little evidence of effects on other observable teacher characteristics.

We next consider two mechanisms by which the CSB policy might improve teacher staffing. First, we restrict the sample to new hires and re-estimate the models with teacher characteristics as the dependent variable. We present plots of the local linear estimates of the relationship between the recruitment outcomes and the forcing variable in **Figure 4** and estimated treatment effects in **Table 6**. We find that the CSB policy increased the probability that a newly hired teacher is an NBCT by a small but statistically significant amount, about 1 percentage point.[25] This effect is robust to the inclusion of additional covariates as well as using the alternative forcing variable. As with the overall teacher staffing outcomes, we do not find consistent evidence that the incentive improved other observable teacher credentials of new hires. Simple regression discontinuity estimates suggest that the policy increased the proportion of new hires with advanced degrees; however, this result is not robust to the inclusion of school covariates and we do not find a similar result when we use the CSRS-based forcing variable.

Second, we limit the sample to National Board Certified Teachers and re-estimate our regression discontinuity models. In contrast with the previous results, these estimates are largely descriptive. Our prior models estimate differences in school-wide outcomes at the eligibility threshold under the assumption that assignment near this point is as good as random. However, individual teachers do manipulate their treatment status by deciding whether or not to apply for National Board certification and we do see evidence that this decision does vary discontinuously at the eligibility threshold. Given this, these results should be interpreted as effects on overall NBCT turnover more generally and not as a causal retention effect for individual teachers. Additionally, the sample of teachers is fairly small and the NBCT turnover effects are not as precisely estimated than the other discontinuity effects. We find some suggestive evidence that the policy may have reduced NBCT turnover, but these results are sensitive to choice of bandwidth and covariates. With covariates, we estimate NBCTs in eligible schools are about 2 percentage points less likely to leave, but this difference is only statistically significant at the 10% level. Point estimates using the alternative forcing variable are similar, but not statistically significant and less precisely estimated. Using the optimal bandwidth for NBCT turnover effects, the estimated coefficient from the model with covariates is similar and statistically significant

[25]The mean number of new hires for schools in our sample is 3.8, which suggests an average effect of approximately 0.04 additional NBCTs hired each year. As we show in Table 3, NBCTs comprise about 8% of our sample as a whole. While this result is not robust to using smaller bandwidths, the optimal bandwidth for NBCT status among new hires is much larger than the one we use here; the estimated coefficients using this bandwidth are nearly identical. See Figure A.2 and Table A.1.

at the 5% level; the estimated coefficient for the simple RD model is never statistically significant.[26]

Taken together, our results suggest that the CSB policy increased the proportion of NBCTs in high poverty schools by about 2-3 percentage points over the first years of implementation. We find that teachers are about 50% more likely to apply for NBPTS certification in eligible schools. We also find an increase in the proportion of new hires that have been previously certified. The policy may have reduced turnover of NBCTs, however we do not find definitive evidence that this occurred. As we discuss in the literature review above, the implications of the staffing effects for student achievement are theoretically unclear and may depend on how participation in the application process influences teacher effectiveness both during the application year and in later years. Given the potential ambiguity about what these results suggest for the quality of teaching in high-poverty schools, we now turn to a direct assessment of the policy effects on student achievement.

IV Teacher Bonuses and Student Achievement

We begin our presentation of the student achievement results with graphical evidence from the regression discontinuity designs in **Figure 5**. In Figures 5a - 5d, we graph the simple RDD without controls for prior student achievement using both the full sample and the pretest sample. The relationship between the forcing variable and test scores appears to be linear and none of the graphs shows a clear discontinuity at the eligibility threshold. In Figures 5e and 5f, we first remove the variation in student achievement explained by the lagged achievement measures and student characteristics and plot the regression residuals against the forcing variable. In both cases, there is some evidence of a negative effect at the threshold.

We investigate the student achievement effects more formally in **Table 7**. Consistent with the graphical evidence, we uncover little evidence of student achievement effects. In Panel A, we present estimates using the simple sharp RD from Eq. (2). We estimate discontinuities in the conditional expectation of student achievement at the eligibility threshold of -0.004 in math and -0.022 in reading, but neither of these is statistically significant. When we include student and school covariates, these estimates are -0.013 and -0.009, respectively, but not statistically significant. The results are similar when we include individual student pretest scores from 2007, which we display in columns (5) and (10). The effect of CSB attendance on achievement is smaller at -0.030 for math and -0.015 for reading, but

[26]See Figure A.2 and Table A.1.

again neither is statistically significant.[27]

The regression discontinuity models provide an estimate of the average treatment effect for students in schools on the margin of eligibility for the CSB. If we model achievement as a cumulative process and allow the effect of attendance at a CSB school to depend on the length of attendance, then the regression discontinuity estimate is a weighted average treatment effect of the effects of differing lengths of exposure, where the weights depend on the likelihood that students have attended a CSB for a given duration. To give these estimates greater meaning, we present estimated treatment effects from the instrumental variables model in Eq. (3) that instruments length of attendance at a CSB school with the eligibility threshold indicator for the current year. This estimator scales the reduced-form treatment effect in Panel A by the discontinuity at the eligibility threshold in the length of treatment to generate an estimate of the effect of an additional year of attendance at a CSB school. Point estimates suggest an effect of one year of attendance of between 0 and about -0.02 standard deviations, which are similar to the OLS estimates in columns (1) and (6), but as with the treatment effects in Panel A, these results are statistically insignificant.

The school staffing results in Section III indicate that eligibility for the CSB improves the observable qualifications of teachers. The lack of positive student achievement effects may therefore be somewhat surprising. However, given the imprecision in the estimates even with student pretest scores, we cannot rule out small positive effects. For instance, research in Washington suggests NBCTs are about 0.02 standard deviations more effective than the average non-NBCT in elementary classrooms (Cowan and Goldhaber, 2015). By comparison, the upper bounds of the 95% confidence intervals for the regression discontinuity models with covariates extend to about 0.01-0.02 standard deviations. The estimated staffing effects may therefore not be large enough to generate substantial student achievement effects. Instrumental variables models using the matched student-teacher data suggest that an additional year of CSB eligibility increases the likelihood that a teacher in a tested grade is an NBCT by about 0.024 in math and 0.016 in reading. These results suggest an increase in the number of NBCTs by 2013 of about 0.15 and 0.10, respectively, in schools eligible in each year. Given these impacts, the available evidence on the effectiveness of NBCTs suggests average improvements in teacher quality of less than 0.01 student standard deviations, which is not inconsistent with our findings.

[27]We show the estimated treatment effects both with and without covariates for different bandwidths in Appendix Figure 3. In both cases, the estimated treatment effects are relatively stable with respect to the choice of bandwidth.

Additional Robustness Checks

In Section II, we provided some evidence of violations of the RDD assumptions. Despite evidence of manipulation of the forcing variable, the results presented thus far are robust to inclusion of controls for pre-treatment outcomes and the use of an alternative forcing variable. An additional concern with the longitudinal nature of the student data is that school choice may be related to eligibility for the CSB program. School choice could bias our results in at least two ways. Staffing improvements at eligible schools may lead to improvements in student retention and differentially affect the composition of test takers in eligible schools near the threshold. It is difficult to assess the direction of this potential bias a priori. NBCTs tend to teach students with higher achievement in Washington and programming for advanced students may influence school choice decisions (Cowan and Goldhaber, 2015; Epple et al., 2002). In this case, we may expect that eligible schools differentially retain high-achieving students, which would bias upward our treatment effect estimates. On the other hand, schools with high mobility tend to serve disadvantaged and low-achieving students (Xu et al., 2009). To the extent that school staffing improves the retention of low-achieving students, estimated treatment effects may be biased downward. Even if students in eligible schools do not attrit at different rates, school choice may also bias our results if students re-sort along observable or unobservable dimensions into eligible schools after the introduction of the policy. School choice may bias the results either through student transfers or by selection into eligible schools at promotional milestones.

We have already presented some evidence that school choice does not bias our results in Section II. When we test for differences in 2007 baseline achievement during the treatment period at the eligibility threshold, we find no statistically significant differences. This suggests that any attrition differences are not related to baseline student outcomes. We now expand upon this preliminary test for selective school choice in two ways. First, we directly test for differential attrition at the eligibility threshold by re-estimating our regression discontinuity models using an indicator for student attrition as the dependent variable. We define a student as attriting if he does not have a test score in the following year but is expected to given the state's testing schedule.[28] We present these results in **Table 8**. We uncover little evidence of selective attrition. Point estimates in math and reading range from -0.9 to 0.2 percentage points and are not statistically significant. Inclusion of covariates also does little to change the results. It therefore does not appear that differential attrition is important in this context.

We next recast the regression discontinuity design as an intent-to-treat (ITT) analysis

[28]For 8th grade students, we define attrition as not having a 10th grade test score after 2 years.

using the eligibility of the schools that students attended in 2007. For each year t, we use the year-t forcing variable for a students' school of attendance in 2007 to instrument for actual year-t eligibility. Because this analysis only relies on variation in eligibility associated with a students' school in 2007, we circumvent problems associated with non-random re-allocation of students to schools following the introduction of the CSB policy. We conduct the ITT analysis using the sample of students with 2007 testing and school assignment records and present the estimated effects in Table 8. The estimates are largely consistent with the estimates from the pre-test sample in Table 7. The estimated ITT effects are somewhat larger in math and somewhat smaller in reading, but are not statistically significant in either case.

V Conclusion

We study the introduction of an incentive for National Board Certified Teachers to work in high-poverty schools. We find that the bonus increased the proportion of teachers with the NBPTS credential both by incentivizing incumbent teachers to apply for certification and through better recruitment of teachers who already possess the NBPTS credential. We find suggestive evidence that eligible schools have higher retention rates among NBCTs; however, these results are sensitive to modeling decisions. Importantly, the increase in the number of NBCTs does not appear to have increased student achievement. Eligible schools have similar achievement levels as ineligible schools and 95% confidence intervals rule out cumulative effects on student achievement larger than 0.01-0.02 standard deviations.

There are two important limitations of our study. First is the limited time period over which we can assess student achievement outcomes. Depending on the model, we consider student outcomes for the first 5-6 years of implementation. The teacher staffing outcomes suggest that the CSB did increase the observable credentials of teachers in high-poverty schools. While the policy may affect the incentives to apply for NBPTS certification, and hence the distribution of teacher quality among NBCTs, Cowan and Goldhaber (2015) find no evidence that NBCTs in CSB-eligible schools are less effective relative to their peers than NBCTs in ineligible schools. As we find some suggestive evidence that increases in applications may counteract achievement effects due to improved staffing, it is possible that reductions in student achievement in the short-term may be offset by longer-term improvements in school quality. Second, our assessment is limited to student achievement on standardized tests. While research has found that school and teacher interventions that improve measured student learning also produce long-term effects on educational attainment, criminal behavior, and wages, there may be effects of the CSB policy that we cannot capture with test-based instruments alone (Billings et al., 2013; Chetty et al., 2014b; Deming et al., 2014).

References

Billings, S. B., Deming, D. J., and Rockoff, J. (2013). School segregation, educational attainment, and crime: Evidence from the end of busing in Charlotte-Mecklenburg. *The Quarterly Journal of Economics*, 129(1):435–476.

Borman, G. D. and Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, 78(3):367–409.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., and Wyckoff, J. (2008). Who leaves? Teacher attrition and student achievement. Technical report, National Bureau of Economic Research, Cambridge, MA.

Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2013). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers. *Journal of Labor Economics*, 31(1):83–117.

Boyd, W. L. and Reese, J. P. (2006). Great expectations: The impact of the National Board for Professional Teaching Standards. *Education Next*, 6(2):50–57.

Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *The Journal of Political Economy*, 100(1):1–40.

Cascio, E., Gordon, N., Lewis, E., and Reber, S. (2010). Paying for progress: Conditional grants and the desegregation of southern schools. *Quarterly Journal of Economics*, 125(1):445–482.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9):2633–2679.

Chingos, M. M. and Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, 30(3):449–465.

Clotfelter, C. T., Glennie, E., Ladd, H., and Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, 92(5-6):1352–1370.

Clotfelter, C. T., Ladd, H., and Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6):673–682.

Clotfelter, C. T., Ladd, H., and Vigdor, J. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3):655–681.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2011). Teacher mobility, school segregation, and pay-based policies to level the playing field. *Education Finance and Policy*, 6(3):399–438.

Cowan, J. and Goldhaber, D. (2015). National Board certification and teacher effectiveness: Evidence from Washington. Technical Report 2015-1, Center for Education Data and Research, Seattle, WA.

Dee, T. and Wyckoff, J. (2013). Incentives, selection, and teacher performance: Evidence from IMPACT. Technical Report 19529, National Bureau of Economic Research, Cambridge, MA.

Deming, D. J., Hastings, J. S., Kane, T. J., and Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review*, 104(3):991–1013.

Ehlert, M., Koedel, C., Parsons, E., and Podgursky, M. J. (2013). The sensitivity of value-added estimates to specification adjustments: Evidence from school-and teacher-level models in Missouri. *Statistics and Public Policy*, 1(1):19–27.

Elfers, A. M. and Plecki, M. L. (2014). Results of a state incentive program on the supply and distribution of National Board Certified Teachers. *Leadership and Policy in Schools*, 13(2):147–168.

Epple, D., Newlon, E., and Romano, R. (2002). Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, 83(1):1–48.

Figlio, D. N. (2002). Can public schools buy better-qualified teachers? *Industrial and Labor Relations Review*, 55(4):686–699.

Fulbeck, E. S. (2014). Teacher mobility and financial incentives: A descriptive analysis of Denver's ProComp. *Educational Evaluation and Policy Analysis*, 36(1):67–82.

Glazerman, S., Protik, A., Teh, B., Bruch, J., and Max, J. (2013). Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. Technical Report 2014-4003, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Goldhaber, D. and Anthony, E. (2007). Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. *Review of Economics and Statistics*, 89(1):134–150.

Goldhaber, D., Choi, H.-J., and Cramer, L. (2007). A descriptive analysis of the distribution of NBPTS-certified teachers in North Carolina. *Economics of Education Review*, 26(2):160–172.

Goldhaber, D., Gross, B., and Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management*, 30(1):57–87.

Goldhaber, D. and Hansen, M. (2009). National Board certification and teachers' career paths: Does NBPTS certification influence how long teachers remain in the profession and where they teach? *Education Finance and Policy*, 4(3):229–262.

Goldhaber, D., Lavery, L., and Theobald, R. (2014). Uneven playing field? Assessing the inequity of teacher characteristics and measured performance across students. Technical Report 2014-4, Center for Education Data and Research, Seattle, WA.

Hahn, J., Todd, P. E., and Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Hanushek, E. A., Kain, J. F., O'Brien, D., and Rivkin, S. G. (2005). The market for teacher quality. Technical report, National Bureau of Economic Research, Cambridge, MA.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (1999). Do higher salaries buy better teachers? Technical Report 7082, National Bureau of Economic Research, Cambridge, MA.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Why public schools lose teachers. *The Journal of Human Resources*, 39(2):326–354.

Harris, D. N. and Sass, T. R. (2009). The effects of NBPTS-certified teachers on student achievement. *Journal of Policy Analysis and Management*, 28(1):55–80.

Humphrey, D. C., Koppich, J. E., and Hough, H. J. (2005). Sharing the wealth: National Board certified teachers and the students who need them most. *Education Policy Analysis Archives*, 13(18).

Imazeki, J. (2005). Teacher salaries and teacher attrition. *Economics of Education Review*, 24(4):431–449.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959.

Jackson, C. K. and Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4):85–109.

Jacob, B. A. (2007). The challenges of staffing urban schools with effective teachers. *The Future of Children*, 17(1):129–153.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1):37.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Loeb, H., Elfers, A. M., Plecki, M. L., Ford, B., and Knapp, M. S. (2006). National Board certified teachers in Washington State: Impact on professional practice and leadership opportunities. Technical report, Center for Strengthening the Teaching Profession, Seattle, WA.

Loeb, S. and Page, M. E. (2000). Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary variation. *Review of Economics and Statistics*, 82(3):393–408.

Matsudaira, J. D., Hosek, A., and Walsh, E. (2012). An integrated assessment of the effects of Title I on school behavior, resources, and student achievement. *Economics of Education Review*, 31(3):1–14.

McCrary, J. (2008a). DCdensity.

McCrary, J. (2008b). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(698-714):1–17.

National Board for Professional Teaching Standards (2010). Profiles in excellence: Washington State. Technical report, National Board for Professional Teaching Standards, Arlington, VA.

National Board for Professional Teaching Standards (2014a). 2014 state rankings by new number of National Board certified teachers. Technical report, National Board for Professional Teaching Standards, Arlington, VA.

National Board for Professional Teaching Standards (2014b). 2014 state rankings by total number of National Board certified teachers. Technical report, National Board for Professional Teaching Standards, Arlington, VA.

National Center for Education Statistics (2015). Schools and Staffing Survey, Public School Teachers: 2011-2012. Technical report, National Center for Education Statistics, Washington, D.C.

Plecki, M. L., Elfers, A. M., St. John, E., Finster, M., Emry, T., Nishida, N., and Harmon, J. (2010). Study of the incentive program for washington's national board certified teachers. Technical report, The Center for the Study of Teaching and Policy, University of Washington, Seattle, WA.

Podgursky, M. J. and Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management*, 26(4):909–950.

Raudenbush, S. W., Rowan, B., and Cheong, Y. F. (1992). Contextual effects of the self-perceived efficacy of high school teachers. *Sociology of Education*, 65(2):150–167.

Ronfeldt, M., Loeb, S., and Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1):4–36.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., and Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2-3):104–122.

Sato, M., Wei, R. C., and Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board certification. *American Educational Research Journal*, 45(3):669–700.

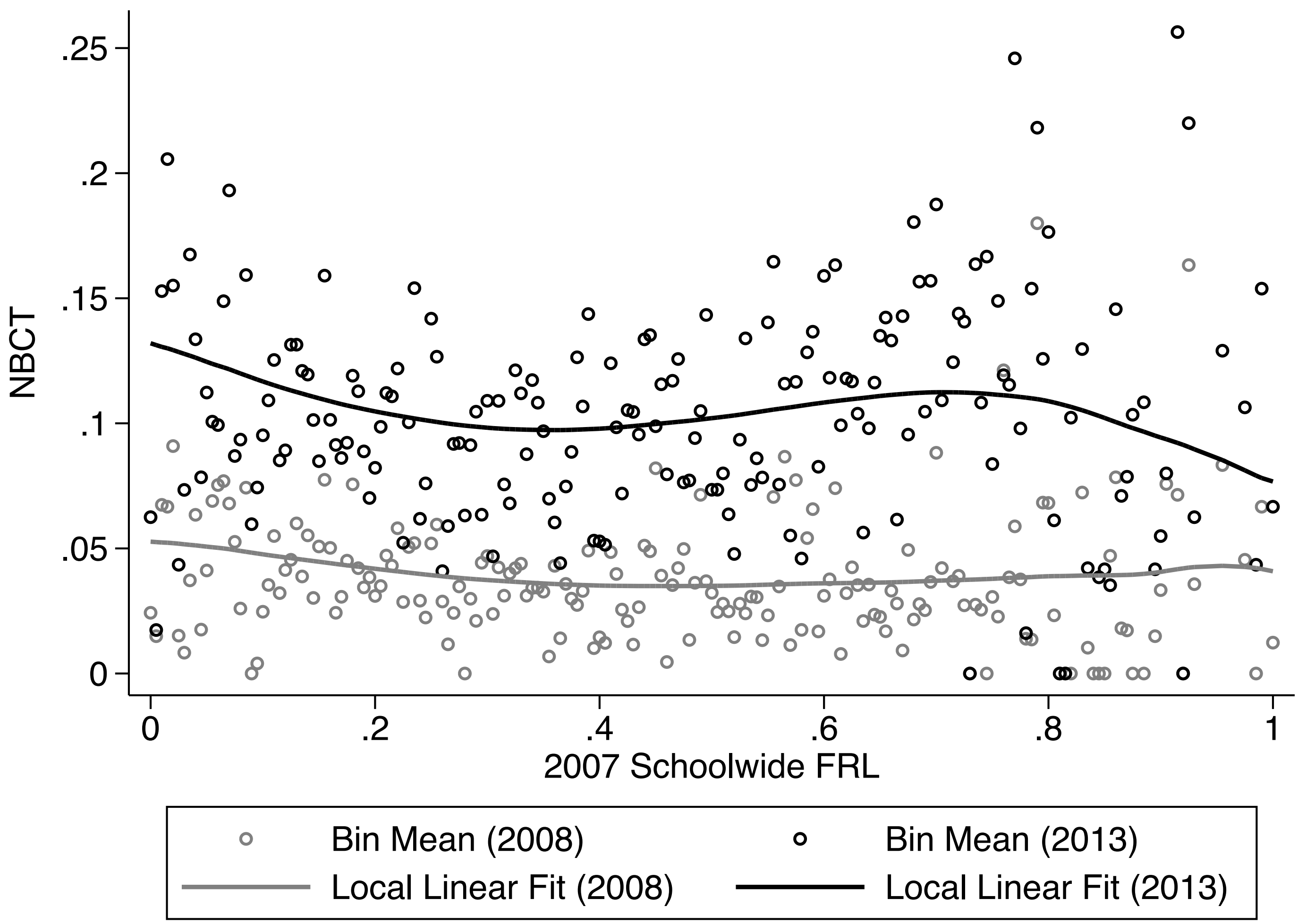
Simpkins, J. (2011). What does washington state get for its investment in bonuses for Board certified teachers? Technical report, Center on Reinventing Public Education, Seattle, WA.

Springer, M. G., Rodriguez, L. A., and Swain, W. A. (2014). Effective teacher retention bonuses: Evidence from Tennessee. Technical report, Tennessee Consortium on Research, Evaluation, and Development, Nashville, TN.

Steele, J. L., Murnane, R. J., and Willett, J. B. (2010). Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. *Journal of Policy Analysis and Management*, 29(3):451–478.

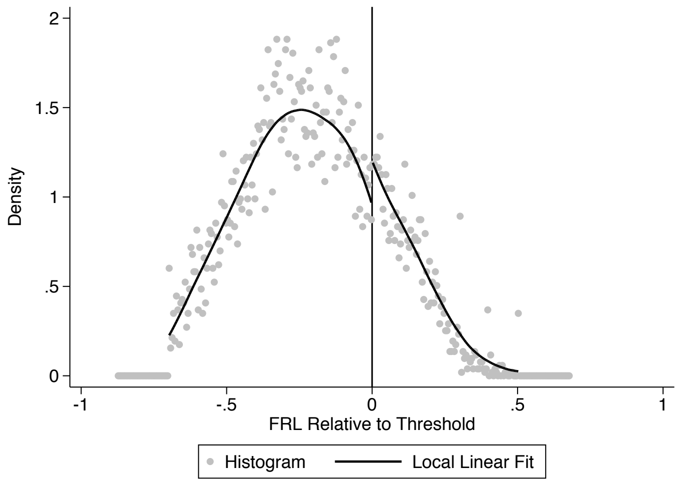
Xu, Z., Hannaway, J., and D'Souza, S. (2009). Student transience in North Carolina: The effect of school mobility on student outcomes using longitudinal data. Technical Report 22, Center for Analysis of Longitudinal Data in Education Research, Washington, D.C.

Figure 1: NBCTs and School Poverty in 2008 and 2013

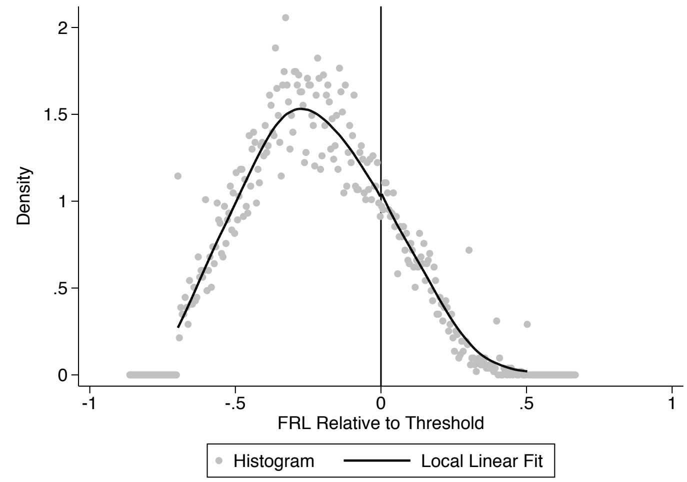


Notes: Figure plots proportion of teachers who are NBCTs by 2006-2007 school FRL enrollment share in 2008 and 2013. The scatter plot contains means for bins of width 0.005 and the lines are local linear plots using a triangular kernel and bandwidth of 0.35.

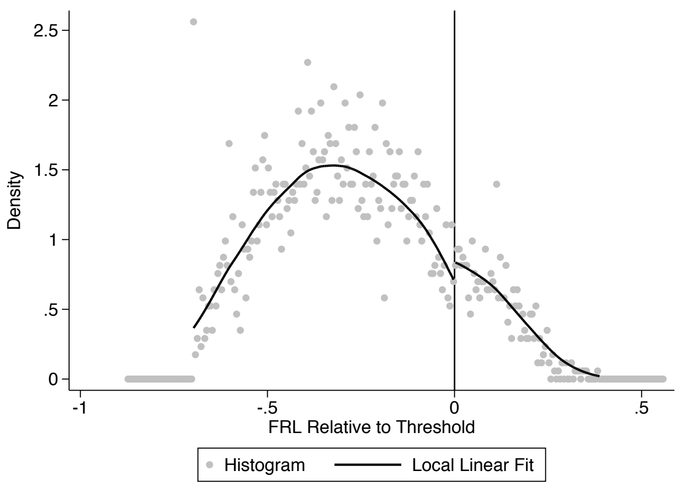
Figure 2: Discontinuity in Forcing Variable Density at Threshold



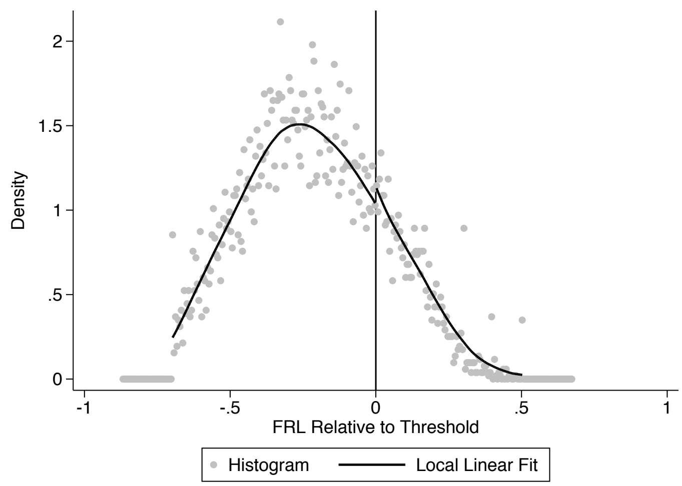
(a) Density of Forcing Variable



(b) Density of CSRS Poverty Measure

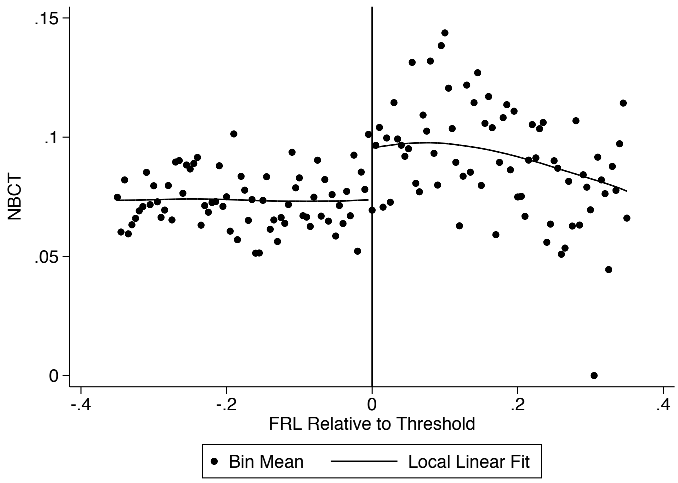


(c) Density of CNS Poverty Measure

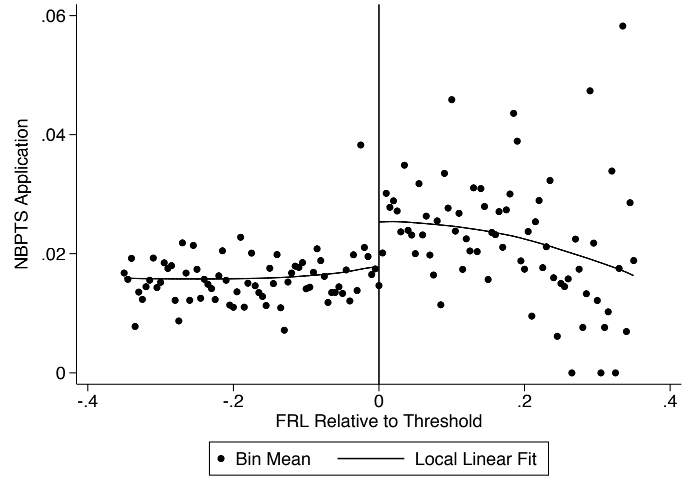


(d) Density of Alternative Forcing Variable

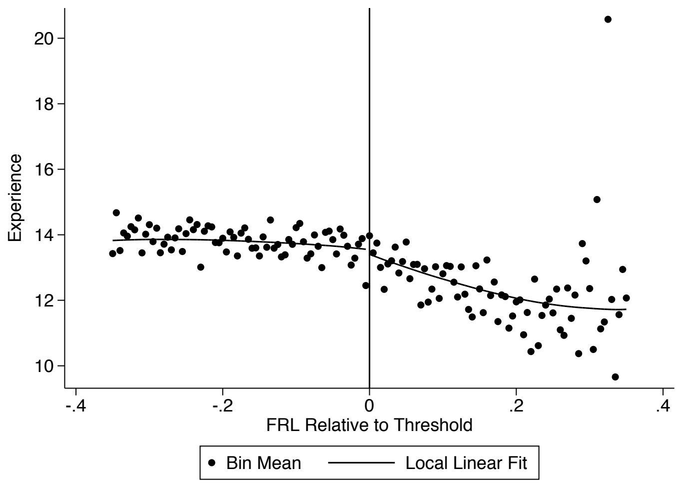
Figure 3: Discontinuity in Teacher Staffing Outcomes at Eligibility Threshold



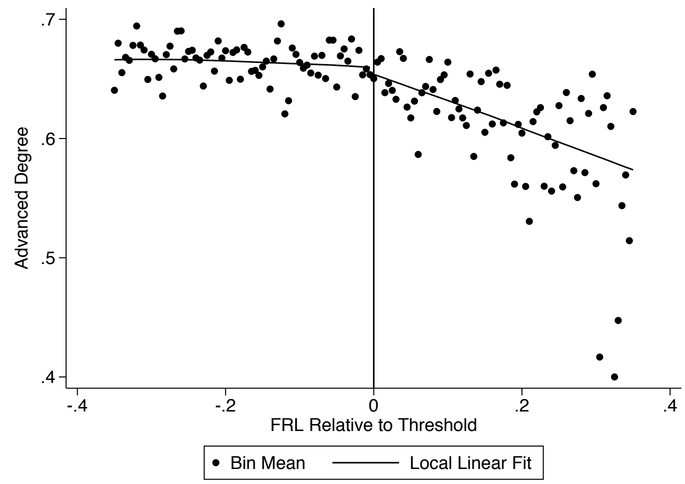
(a) National Board Certified Teacher



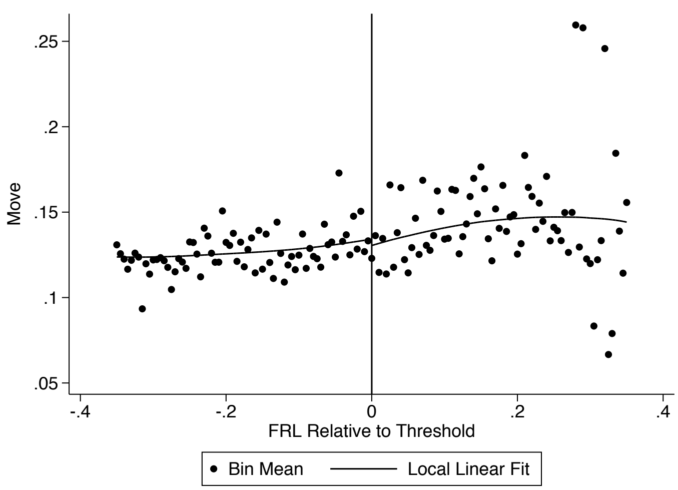
(b) NBPTS Candidate



(c) Experience

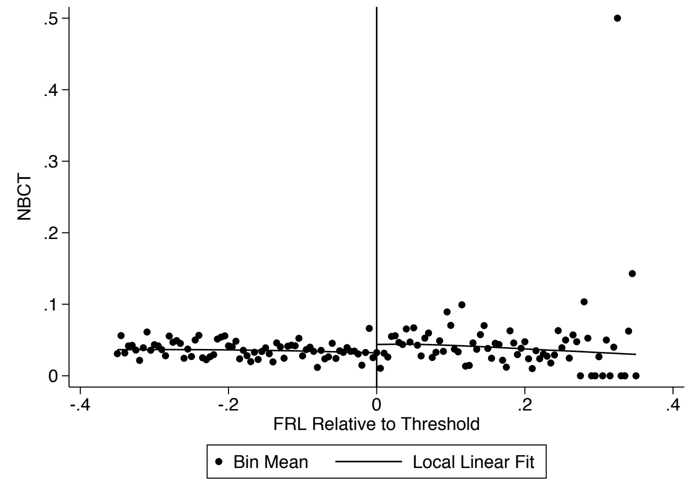


(d) Advanced Degree

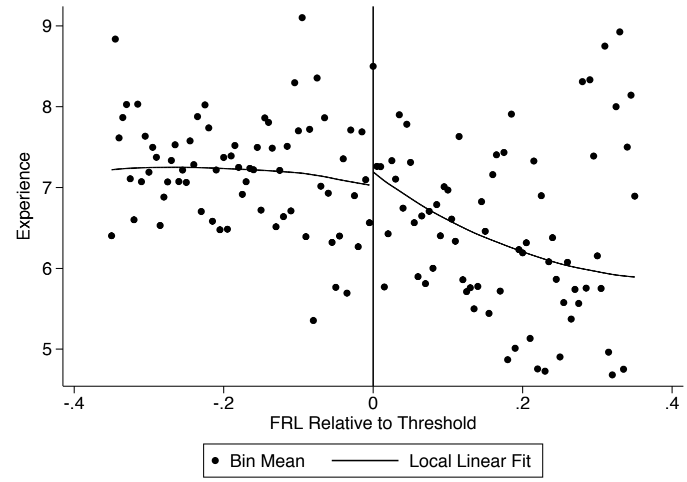


(e) Turnover

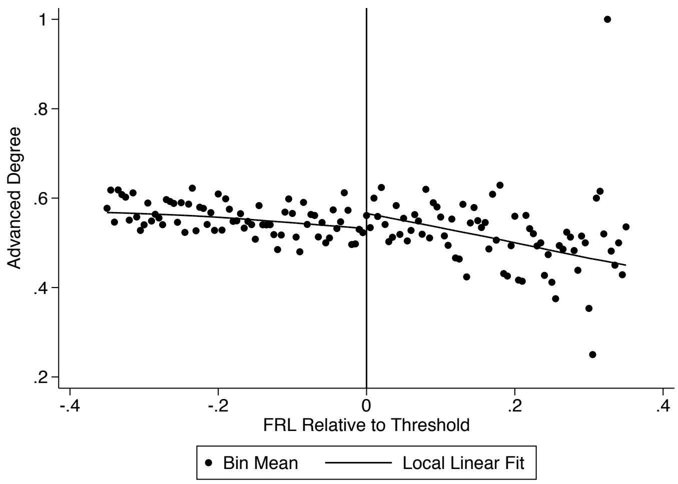
Figure 4: Discontinuity in Teacher Recruitment and Retention Outcomes at Eligibility Threshold



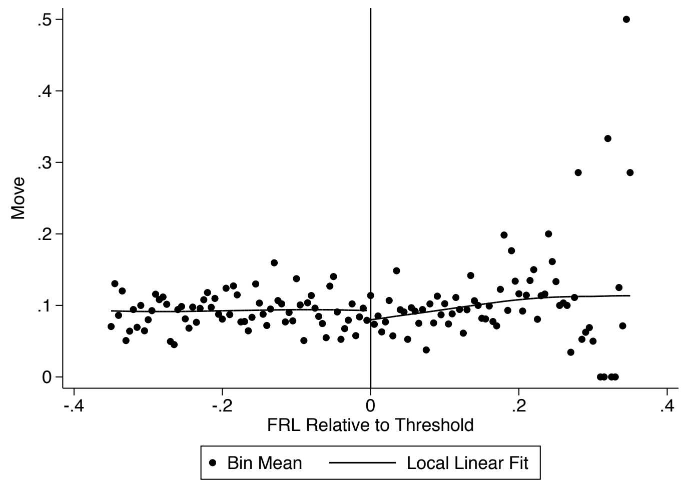
(a) New Hires: NBCT



(b) New Hires: Experience

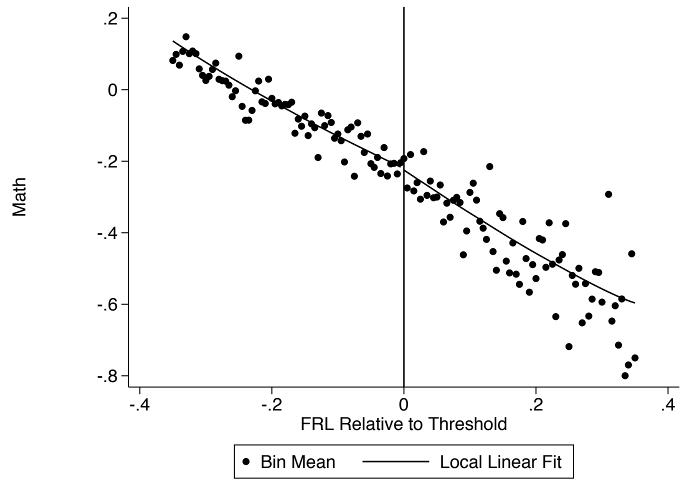


(c) New Hires: Advanced Degree

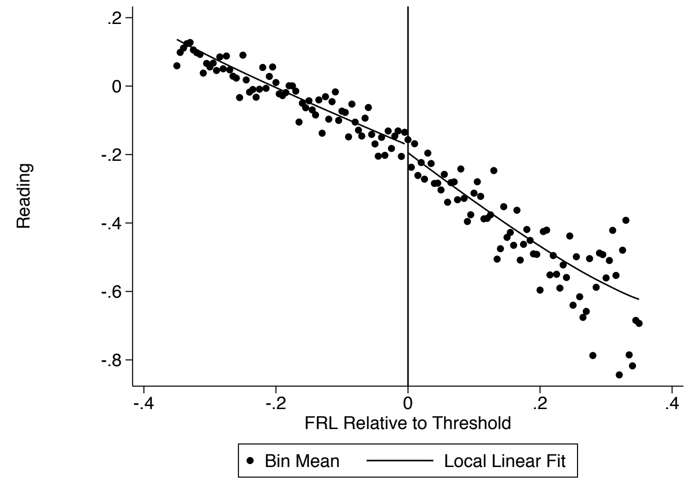


(d) NBCTs: Turnover

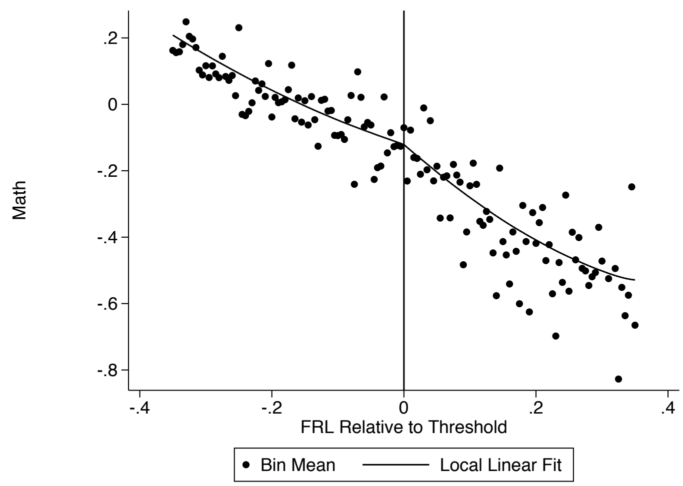
Figure 5: Discontinuity in Student Achievement Outcomes at Eligibility Threshold



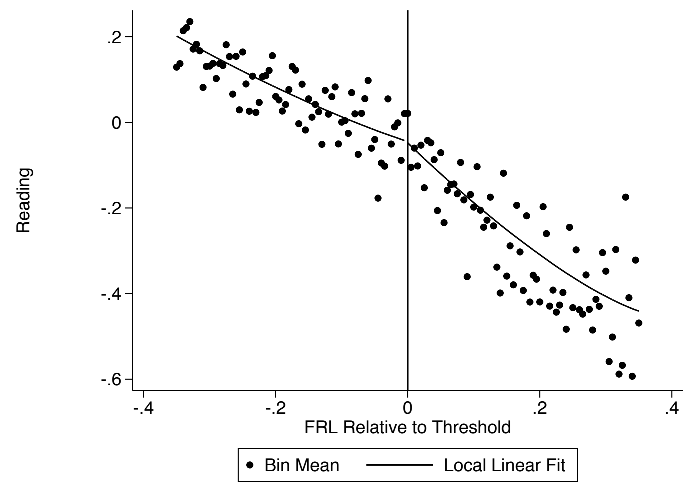
(a) Math (Full Sample)



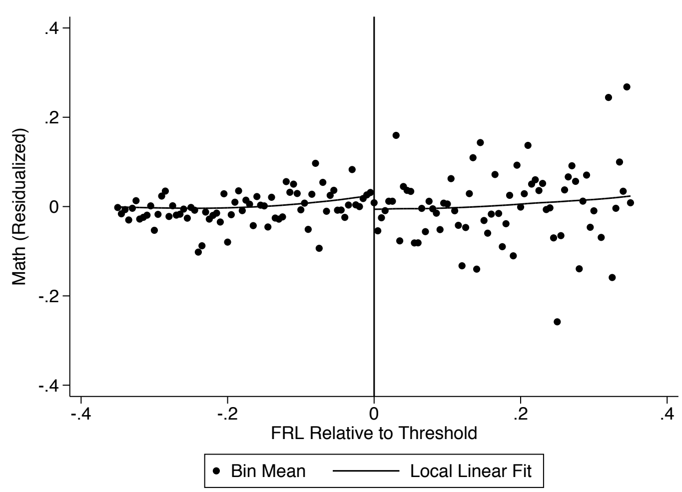
(b) Reading (Full Sample)



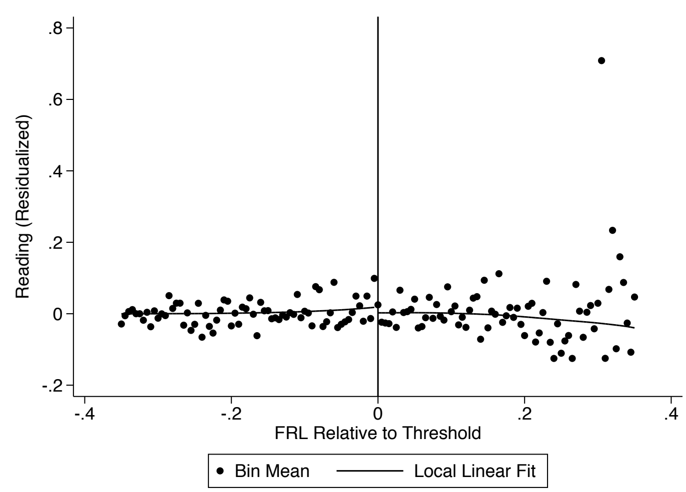
(c) Math (Pretest Sample)



(d) Reading (Pretest Sample)



(e) Math Residuals (Pretest Sample)



(f) Reading Residuals (Pretest Sample)

Table 1: Challenging Schools Bonus Eligibility

Year	Data Source	Prior Year?	Threshold		
			Elementary Schools	Middle Schools	High Schools
2007-2007	CSRS/CNS		70%	70%	70%
2008-2009	CSRS/CNS	Yes*	70%	60%	50%
2009-2010	CSRS	Yes	70%	60%	50%
2010-2011	CSRS	Yes	70%	60%	50%
2011-2012	CSRS	No	70%	60%	50%
2012-2013	CSRS	No	70%	60%	50%
2013-2014	CSRS	No	70%	60%	50%

Notes: Eligibility for Challenging Schools Bonus by school level and year. Elementary schools are defined as those with highest grade of 6th or lower. Middle schools are defined as those with a highest grade of 7th-9th grades. High schools are defined as those with a highest grade of 10th-12th. Data source denotes the data series used to estimate FRL enrollments. CSRS = Core Student Record System. In 2009-2010, the Comprehensive Education Data and Research System (CEDARS) replaced the CSRS; we maintain the labeling for simplicity. CNS = Child Nutrition Services report. Prior year indicates schools receiving the Challenging Schools Bonus in a prior year were grandfathered during the current school year. Schools serving fewer than 30 students are also excluded from eligibility unless they are the largest school at that grade level in the district.

*To qualify for the Challenging Schools Bonus, school must have had 70% FRL enrollment in prior year regardless of grade level.

Table 2: Summary Statistics (Teachers)

	Full Sample		Matched Sample			
			Math		Reading	
	All	CSB	All	CSB	All	CSB
Leave school	0.129	0.142	0.110	0.127	0.110	0.127
	(0.335)	(0.349)	(0.313)	(0.333)	(0.313)	(0.333)
NBCT	0.078	0.095	0.106	0.131	0.115	0.138
	(0.268)	(0.293)	(0.308)	(0.337)	(0.319)	(0.345)
NBCT candidate	0.018	0.024	0.018	0.026	0.019	0.023
	(0.132)	(0.152)	(0.133)	(0.160)	(0.138)	(0.151)
Salary	60504.14	60567.28	62992.63	61780.64	62784.80	61912.87
	(15593.08)	(15377.26)	(13718.41)	(14034.08)	(13586.83)	(14119.62)
Experience	13.487	12.517	13.206	11.339	13.285	11.696
	(9.612)	(9.553)	(9.371)	(8.858)	(9.252)	(9.028)
Advanced degree	0.656	0.625	0.697	0.661	0.704	0.692
	(0.475)	(0.484)	(0.459)	(0.473)	(0.456)	(0.462)
New to school	0.128	0.143	0.107	0.131	0.102	0.121
	(0.335)	(0.350)	(0.309)	(0.337)	(0.302)	(0.326)
N	298267	62635	1084366	216428	1019320	202821

Notes: Summary statistics for sample of teachers and matched student-teacher sample. Summary statistics in columns (1) and (2) describe the sample of teachers for all schools and CSB-eligible schools, respectively, with observations at the teacher-year level. Summary statistics in columns (3) - (6) describe the characteristics of teachers in the matched student-teacher sample. Observations are at the student-year level.

Table 3: Summary Statistics (Students)

	Math				Reading			
	Full Sample		Matched Sample		Full Sample		Matched Sample	
	All	CSB	All	CSB	All	CSB	All	CSB
Math test	0.013	-0.375	0.084	-0.264	0.023	-0.366	0.089	-0.259
	(0.997)	(0.944)	(0.981)	(0.932)	(0.995)	(0.942)	(0.979)	(0.930)
Reading test	0.048	-0.324	0.107	-0.236	0.008	-0.373	0.075	-0.276
	(0.975)	(0.970)	(0.948)	(0.948)	(0.997)	(0.990)	(0.966)	(0.962)
Male student	0.502	0.501	0.497	0.496	0.503	0.503	0.497	0.496
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Bilingual student	0.051	0.148	0.044	0.122	0.062	0.170	0.053	0.144
	(0.221)	(0.355)	(0.204)	(0.327)	(0.241)	(0.376)	(0.225)	(0.351)
Learning Assistance Program	0.067	0.128	0.065	0.121	0.066	0.124	0.066	0.121
	(0.250)	(0.334)	(0.247)	(0.326)	(0.248)	(0.330)	(0.248)	(0.326)
Special education student	0.064	0.068	0.046	0.049	0.074	0.079	0.057	0.059
	(0.245)	(0.252)	(0.210)	(0.215)	(0.262)	(0.269)	(0.231)	(0.236)
Gifted student	0.042	0.025	0.045	0.029	0.040	0.023	0.044	0.028
	(0.200)	(0.156)	(0.207)	(0.167)	(0.195)	(0.150)	(0.205)	(0.164)
Asian	0.087	0.080	0.090	0.085	0.087	0.081	0.090	0.085
	(0.282)	(0.271)	(0.286)	(0.278)	(0.283)	(0.273)	(0.286)	(0.279)
Black	0.050	0.082	0.045	0.076	0.050	0.081	0.045	0.075
	(0.217)	(0.275)	(0.207)	(0.264)	(0.217)	(0.273)	(0.208)	(0.264)
Hispanic	0.171	0.424	0.173	0.409	0.177	0.429	0.179	0.418
	(0.377)	(0.494)	(0.378)	(0.492)	(0.382)	(0.495)	(0.383)	(0.493)
White	0.631	0.347	0.626	0.362	0.624	0.343	0.620	0.354
	(0.483)	(0.476)	(0.484)	(0.480)	(0.484)	(0.475)	(0.485)	(0.478)
N	2470049	465035	1124715	228894	2711038	529301	1148299	237525

Notes: Summary statistics for samples of students with math and reading test scores. Pretest sample includes students with 2007 test scores in math and reading. Matched sample includes students linked definitively to teachers. Observations are at the student-year level.

Table 4: Test of Baseline Differences in Student and Teacher Characteristics

	NBCT (1)	Exp. (2)	Adv. Deg. (3)	Turnover (4)	Math (5)	Reading (6)
Panel A. 2007 School Characteristics						
Above Threshold	-0.007**	-0.242	-0.014	0.007	-0.024	-0.026
	(0.003)	(0.234)	(0.010)	(0.007)	(0.020)	(0.018)
N	7123	7123	7123	7123	7123	7123
Above Alternative Threshold	-0.005	-0.151	-0.023**	-0.002	-0.016	-0.010
	(0.003)	(0.224)	(0.009)	(0.007)	(0.018)	(0.017)
N	6939	6939	6939	6939	6939	6939
	Gifted (7)	LEP (8)	SPED (9)	Black (10)	2007 Math (10)	2007 Reading (12)
Panel B. Student Characteristics						
Above Threshold	-0.001	-0.000	0.005*	0.020***	0.030	0.012
	(0.009)	(0.006)	(0.003)	(0.007)	(0.021)	(0.019)
N	1864467	1864467	1864467	1864467	676960	676960
Above Alternative Threshold	-0.000	0.004	0.004	0.021***	0.020	0.017
	(0.009)	(0.006)	(0.003)	(0.007)	(0.021)	(0.019)
N	1815501	1815501	1815501	1815501	650252	650252

Notes: Regression tests for discontinuities in teacher and student characteristics at eligibility threshold. Each cell contains the coefficient on exceeding the eligibility threshold for the given outcome. Panel A contains regression discontinuity models estimated using local linear regression with 2008-2013 data and 2007 school average of given outcome as the dependent variable. Panel B contains estimates using local linear regression with student observations and given characteristic as the dependent variable. Math and reading test scores in Panel B are from 2007. Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Regression Discontinuity Estimates of Staffing Effects

	NBCT		Application		Experience		Adv. Deg.		Turnover	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. CSB Eligibility										
CSB	0.022***	0.025***	0.007***	0.007***	-0.142	0.170	-0.006	-0.003	-0.004	-0.006
	(0.007)	(0.005)	(0.002)	(0.002)	(0.217)	(0.125)	(0.009)	(0.005)	(0.005)	(0.005)
N	209892	209892	209892	209892	209892	209892	209892	209892	209892	209892
Panel B. CSB Duration										
CSB	0.013***	0.015***	0.004***	0.004***	-0.082	0.100	-0.003	-0.002	-0.002	-0.003
	(0.004)	(0.003)	(0.001)	(0.001)	(0.124)	(0.074)	(0.005)	(0.003)	(0.003)	(0.003)
N	209892	209892	209892	209892	209892	209892	209892	209892	209892	209892
Panel C. CSRS Poverty Measure										
CSB	0.026***	0.030***	0.005*	0.005**	-0.058	0.232*	-0.012	-0.002	-0.005	-0.007
	(0.007)	(0.006)	(0.003)	(0.003)	(0.241)	(0.138)	(0.010)	(0.006)	(0.007)	(0.007)
N	204205	204205	204205	204205	204205	204205	204205	204205	204205	204205
Panel D. Matched Sample (Math)										
CSB	0.051***	0.055***	0.013*	0.014**	0.084	0.390	0.021	0.020		
	(0.019)	(0.017)	(0.007)	(0.007)	(0.535)	(0.411)	(0.024)	(0.025)		
N	731191	731191	731191	731191	722330	722330	722330	722330		
Panel E. Matched Sample (Reading)										
CSB	0.035*	0.034**	-0.004	-0.004	-0.604	-0.537	0.038*	0.032		
	(0.020)	(0.017)	(0.005)	(0.005)	(0.454)	(0.366)	(0.022)	(0.020)		
N	677958	677958	677958	677958	670500	670500	670500	670500		
Covariates	N	Y	N	Y	N	Y	N	Y	N	Y

Notes: Regression discontinuity estimates of school staffing effects. Estimates in Panels A C estimate regression discontinuity models on teacher sample. Results in Panels D and E estimates models using student-year observations using matched student-teacher sample. Included covariates are school level indicators, proportion American Indian, Asian, Hispanic, other non-white, FRL, and male enrollment, log enrollment, and 2007 average math and reading test scores, NBCT status, experience, advanced degree, and turnover, as well as year fixed effects. Estimates in Panels A and B use primary (sharp) regression discontinuity forcing variable. Estimates in Panel C use alternative forcing variable constructed from CSRS poverty measure. Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Regression Discontinuity Estimates of Recruitment and Retention Effects

	New Hires						NBCTs	
	NBCT		Experience		Adv. Deg.		Turnover	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. CSB Eligibility								
CSB	0.012**	0.013**	0.171	0.296	0.034**	0.019	-0.013	-0.019*
	(0.006)	(0.005)	(0.285)	(0.273)	(0.014)	(0.013)	(0.012)	(0.011)
N	26865	26865	26865	26865	26865	26865	16779	16779
Panel B. CSB Duration								
CSB	0.007**	0.008**	0.102	0.182	0.020**	0.012	-0.007	-0.011*
	(0.003)	(0.003)	(0.169)	(0.168)	(0.009)	(0.008)	(0.006)	(0.006)
N	26865	26865	26865	26865	26865	26865	16779	16779
Panel C. CSRS Poverty Measure								
CSB	0.012*	0.013*	0.152	0.381	0.016	0.001	-0.009	-0.017
	(0.007)	(0.007)	(0.333)	(0.318)	(0.018)	(0.017)	(0.013)	(0.013)
N	26100	26100	26100	26100	26100	26100	16485	16485
Covariates	N	Y	N	Y	N	Y	N	Y

Notes: Regression discontinuity estimates of school staffing effects. Estimates in Panels A C estimate regression discontinuity models on teacher sample. Results in Panels D and E estimates models using student-year observations using matched student-teacher sample. Included covariates are school level indicators, proportion American Indian, Asian, Hispanic, other non-white, FRL, and male enrollment, log enrollment, and 2007 average math and reading test scores, NBCT status, experience, advanced degree, and turnover, as well as year fixed effects. Estimates in Panels A and B use primary (sharp) regression discontinuity forcing variable. Estimates in Panel C use alternative forcing variable constructed from CSRS poverty measure. Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Regression Discontinuity Estimates of Student Achievement Effects

	Math					Reading				
	OLS	Regression Discontinuity				OLS	Regression Discontinuity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. CSB Eligibility										
CSB	-0.011	-0.004	-0.013	-0.001	-0.030	-0.008	-0.022	-0.009	-0.003	-0.015
	(0.014)	(0.020)	(0.018)	(0.029)	(0.019)	(0.014)	(0.016)	(0.013)	(0.026)	(0.018)
N	587008	1629084	1629084	587008	587008	666673	1837333	1837333	666673	666673
Panel B. CSB Duration										
CSB	-0.007	-0.003	-0.009	-0.000	-0.024	-0.015***	-0.015	-0.006	-0.002	-0.012
	(0.007)	(0.014)	(0.013)	(0.023)	(0.015)	(0.004)	(0.011)	(0.009)	(0.020)	(0.015)
N	587008	1629084	1629084	587008	587008	666673	1837333	1837333	666673	666673
Panel C. CSRS Poverty Measure										
CSB	-0.011	0.008	-0.011	0.011	-0.028	-0.008	-0.012	-0.005	0.018	-0.006
	(0.014)	(0.024)	(0.022)	(0.042)	(0.026)	(0.014)	(0.019)	(0.015)	(0.035)	(0.023)
N	587008	1581775	1581775	560856	560856	666673	1790010	1790010	640604	640604
Covariates	Y	N	Y	N	Y	Y	N	Y	N	Y
Pretest	Y	N	N	N	Y	Y	N	N	N	Y

Notes: Each cell represents estimated coefficient on CSB treatment variable for test score outcome. Coefficients in columns (1) and (6) are from models estimated on sample of students with 2007 test scores in schools within 0.35 of the eligibility threshold. Coefficients in columns (2) - (5) and (7) - (10) are from regression discontinuity models for students within 0.35 of the eligibility threshold and are estimated by local linear models with triangular kernel. Included covariates are gender and race indicators, indicators for participation in gifted, limited English proficiency, special education, Learning Assistance Program, learning disability status, and 2007 school-by-grade average test scores in math and reading, school-year-by-grade fixed effects, as well as school grade level indicators, proportion American Indian, Asian, Hispanic, other non-white, FRL, and male enrollment, log enrollment, and 2007 average math and reading test scores, NBCT status, experience, advanced degree, and turnover. Pretest controls are 2007 test scores in math and reading interacted with grade and school year. Estimates in Panels A and B use primary (sharp) regression discontinuity forcing variable. Estimates in Panel C use alternative forcing variable constructed from CSRS poverty measure. Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

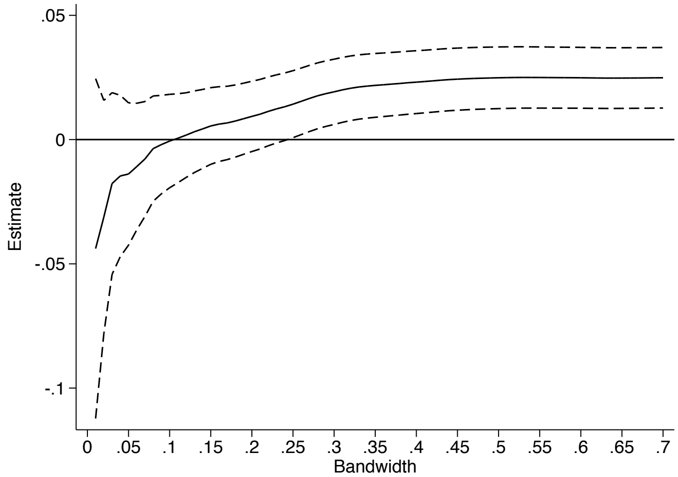
Table 8: Tests for Differential Attrition and School Choice

	Attrition				Intent-to-Treat Effect			
	Math		Reading		Math		Reading	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CSB	0.002	-0.001	-0.009	-0.003	0.045	0.005	-0.011	-0.023
	(0.006)	(0.004)	(0.017)	(0.003)	(0.048)	(0.027)	(0.047)	(0.027)
First stage F-stat.					134.760	221.552	121.967	205.129
Covariates	N	Y	N	Y	N	Y	N	Y
Pretest	N	N	N	N	N	Y	N	Y

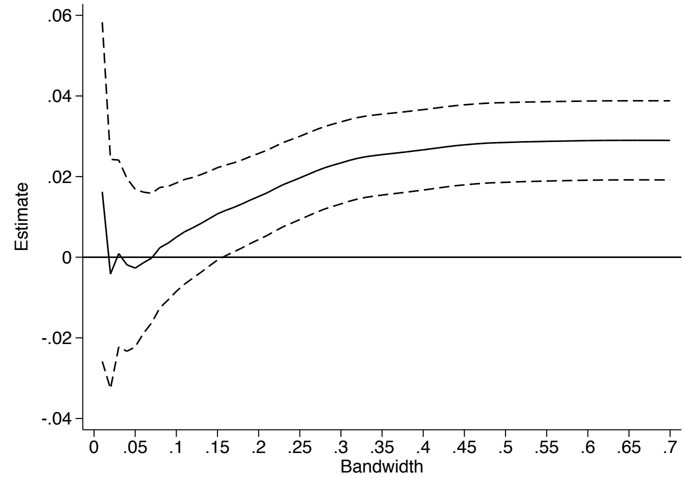
Notes: Each cell represents estimated coefficient on CSB treatment variable. In columns (1)-(4), the dependent variable is the attrition indicator described in the text. In columns (5)-(8), the intent-to-treat analysis instruments student's current eligibility with the current forcing variable and eligibility of the school attended in 2007. The first stage F statistic gives the F statistic on the test of the significance of the coefficient on indicator for exceeding eligibility threshold in the first-stage regression. Included covariates are gender and race indicators, indicators for participation in gifted, limited English proficiency, special education, Learning Assistance Program, learning disability status, and 2007 school-by-grade average test scores in math and reading, school-year-by-grade fixed effects, as well as school grade level indicators, proportion American Indian, Asian, Hispanic, other non-white, FRL, and male enrollment, log enrollment, and 2007 average math and reading test scores, NBCT status, experience, advanced degree, and turnover. Pretest controls are 2007 test scores in math and reading interacted with grade and school year. Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A Bandwidth Choice and Estimation Results

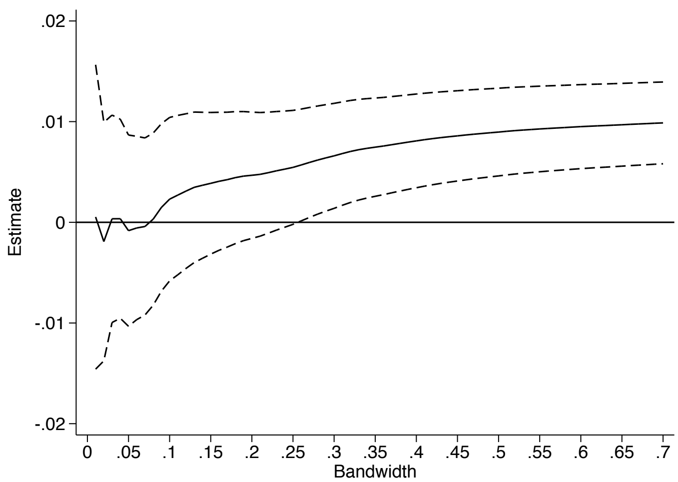
Figure A.1: Sensitivity of Teacher Staffing Outcomes to Bandwidth Choice



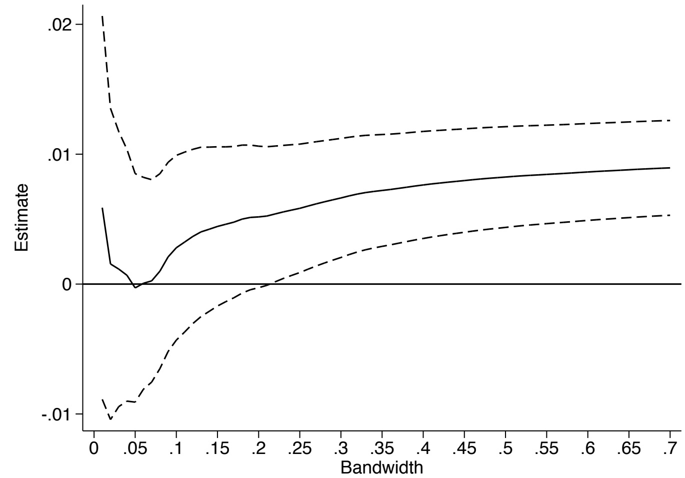
(a) NBCT



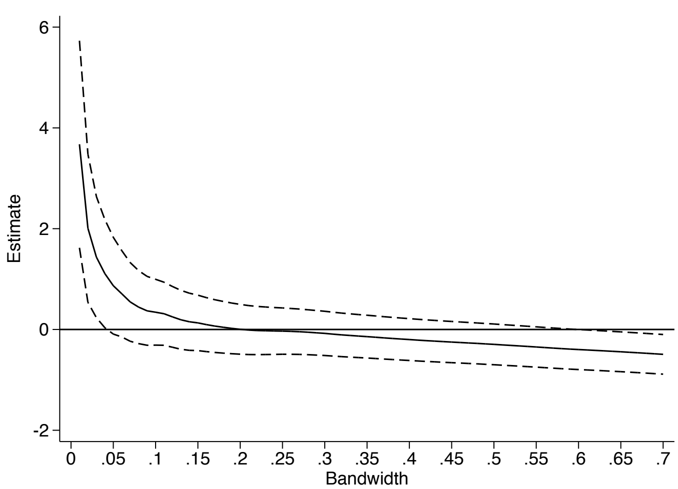
(b) NBCT (w/ covariates)



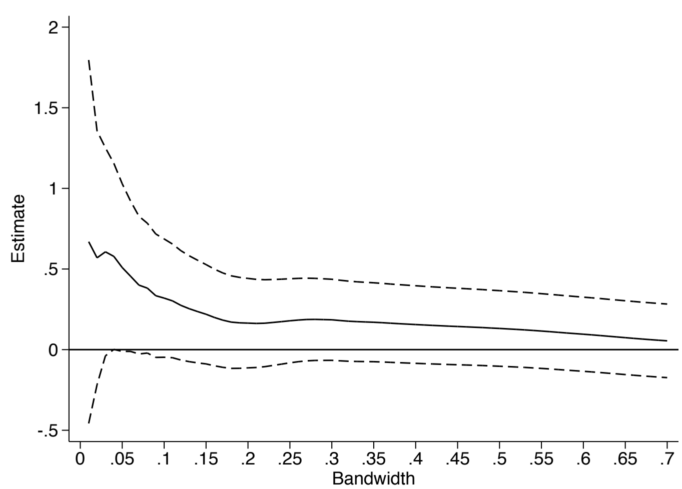
(c) NBPTS Candidate



(d) NBPTS Candidate (w/ covariates)

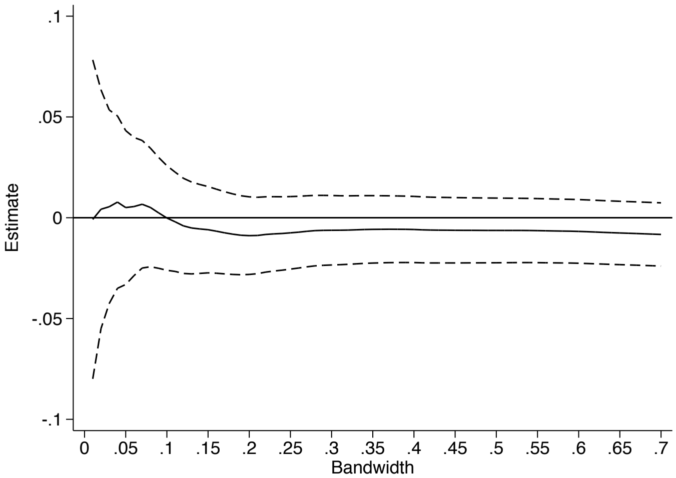


(e) Experience

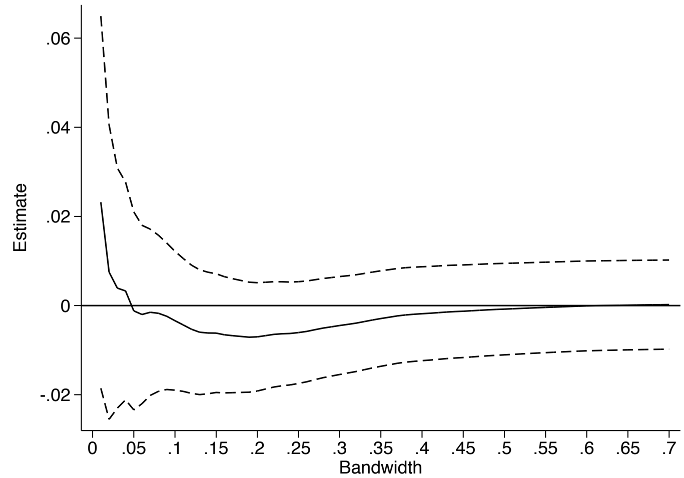


(f) Experience (w/ covariates)

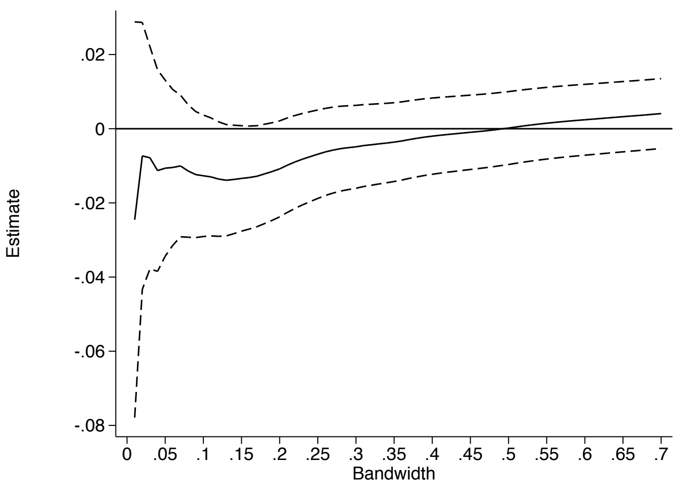
Figure A.1: Sensitivity of Teacher Staffing Outcomes to Bandwidth Choice, cont.



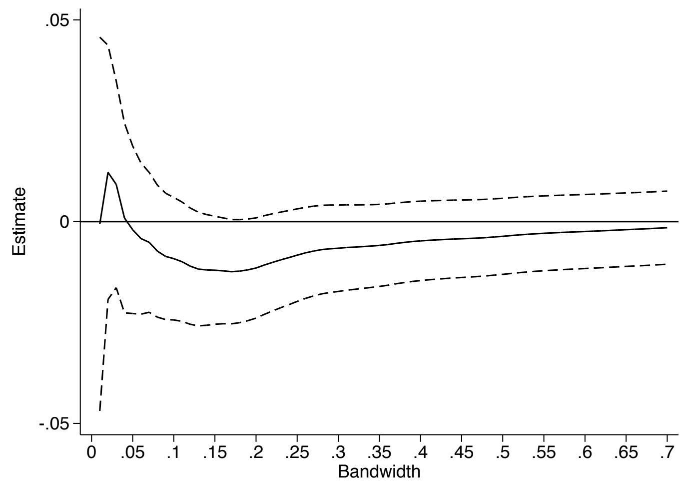
(g) Advanced Degree



(h) Advanced Degree (w/ covariates)



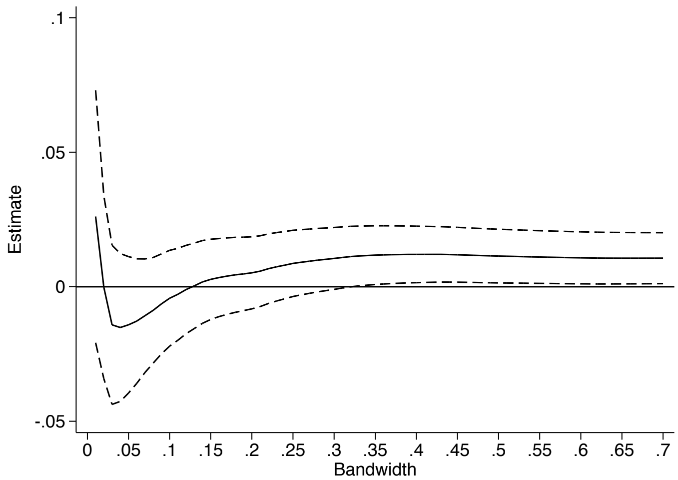
(i) Turnover



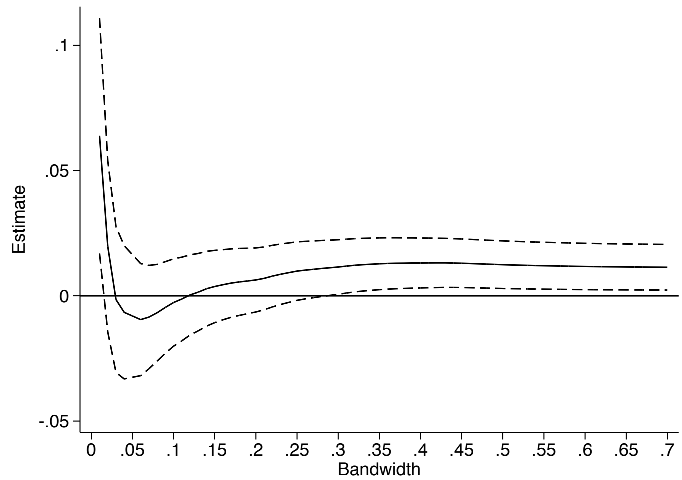
(j) Turnover (w/ covariates)

Notes: Estimated coefficient and 95% confidence interval of treatment effect from local linear regressions with triangular kernel of given bandwidth.

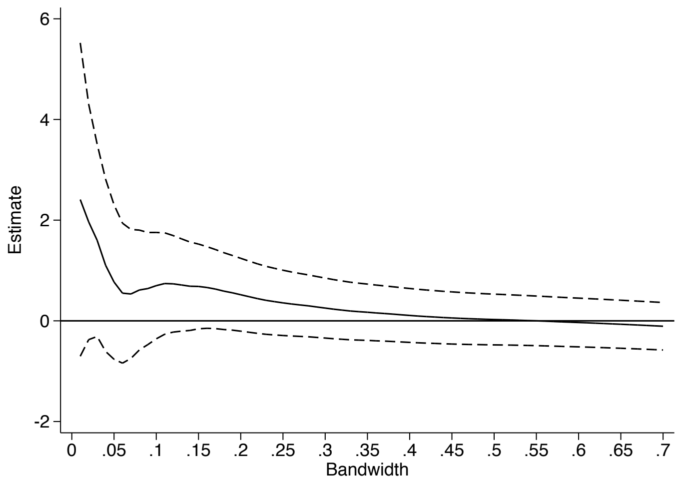
Figure A.2: Sensitivity of Teacher Recruitment and Retention Outcomes to Bandwidth Choice



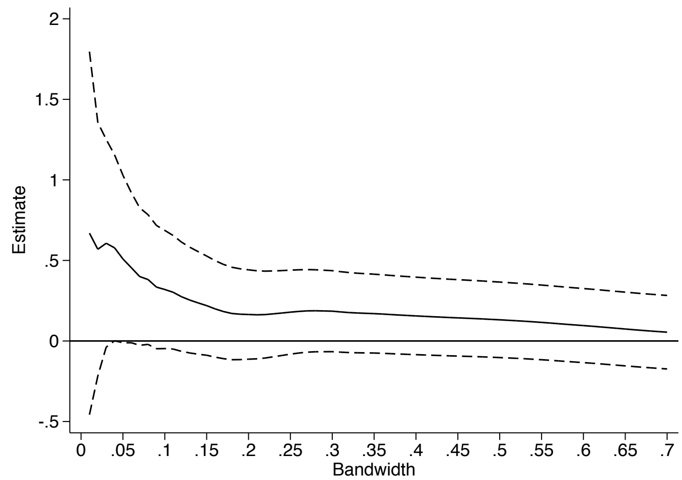
(a) New Hires: NBCT



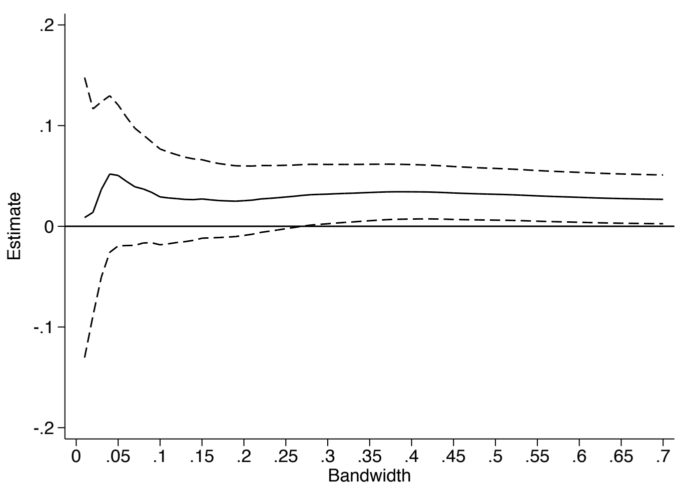
(b) New Hires: NBCT (w/ covariates)



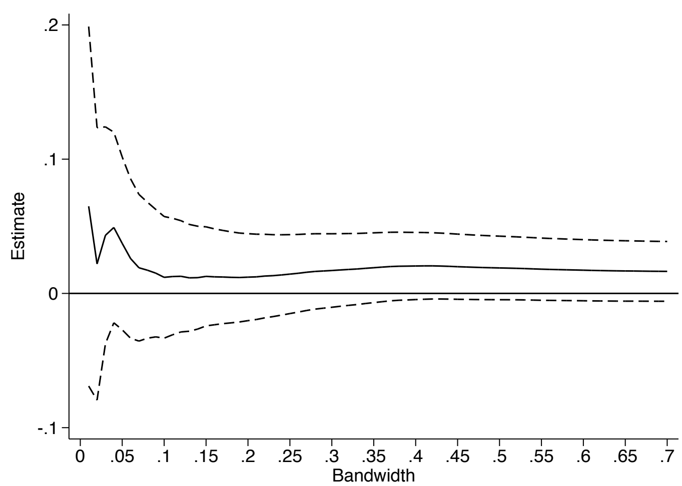
(c) New Hires: Experience



(d) New Hires: Experience (w/ covariates)



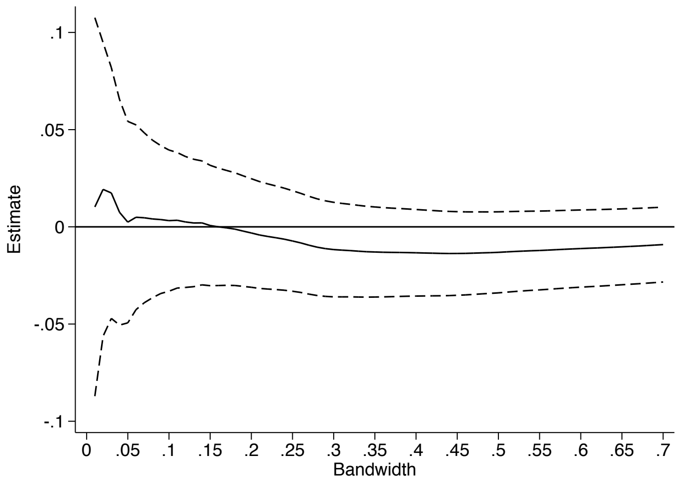
(e) New Hires: Advanced Degree



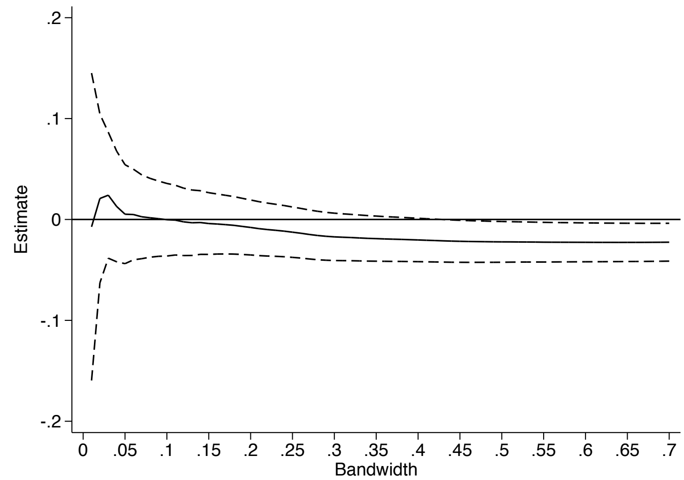
(f) New Hires: Advanced Degree (w/ covariates)

Notes: Estimated coefficient and 95% confidence interval of treatment effect from local linear regressions

Figure A.2: Sensitivity of Teacher Recruitment and Retention Outcomes to Bandwidth Choice, cont.



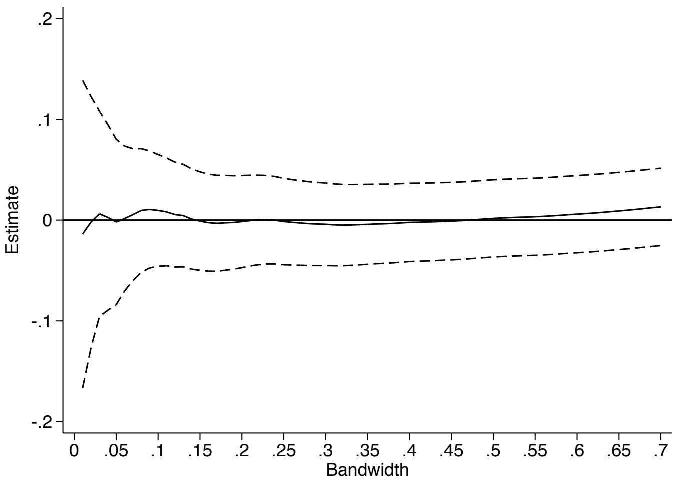
(g) NBCTs: Turnover



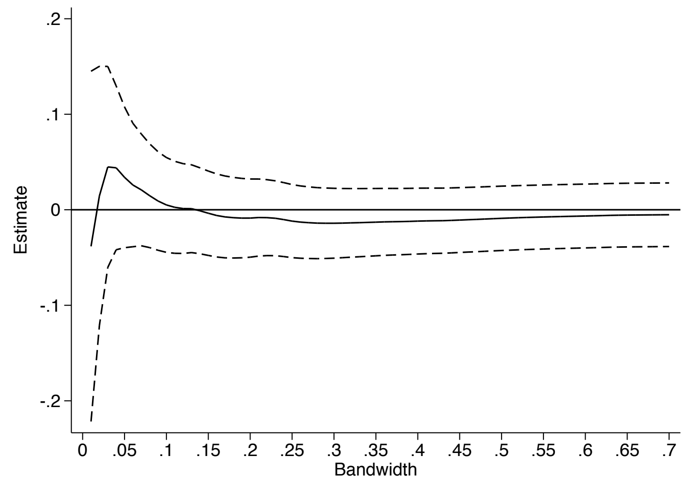
(h) NBCTs: Turnover (w/ covariates)

with triangular kernel of given bandwidth.

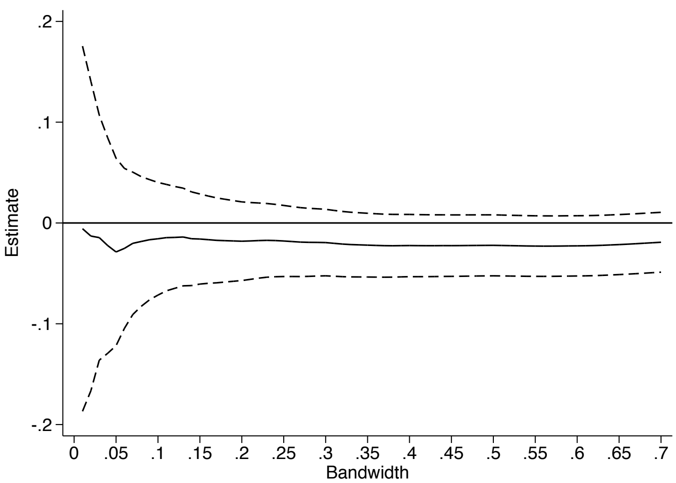
Figure A.3: Sensitivity of Student Achievement Outcomes to Bandwidth Choice



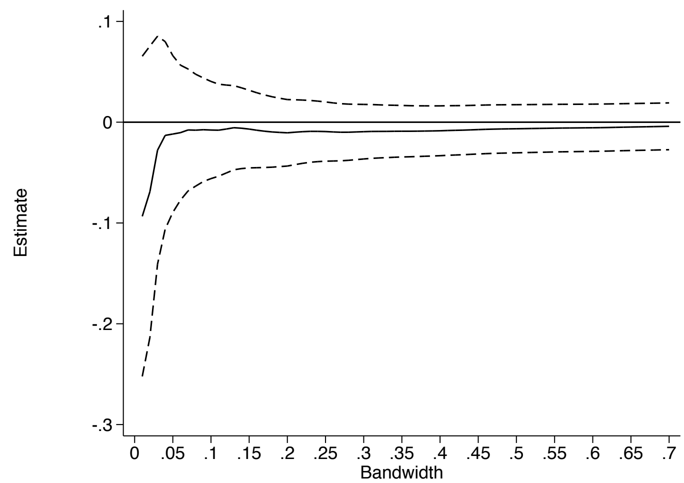
(a) Math (Full Sample)



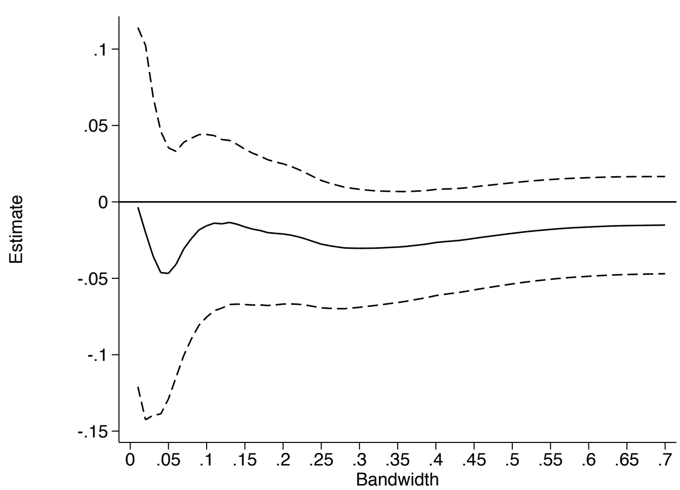
(b) Math (Full Sample w/ covariates)



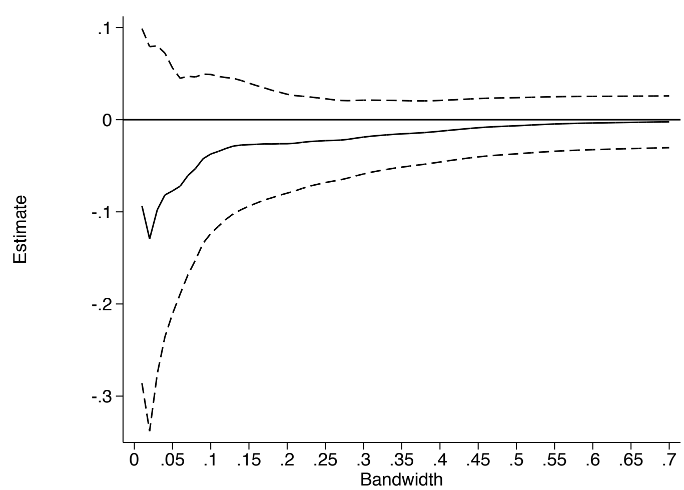
(c) Reading (Full Sample)



(d) Reading (Full Sample w/ covariates)



(e) Math (Pretest Sample w/ covariates)



(f) Reading (Pretest Sample w/ covariates)

Notes: Estimated coefficient and 95% confidence interval of treatment effect from local linear regressions with triangular kernel of given bandwidth.

Table A.3: Estimated Regression Discontinuity Effects with Optimal Bandwidth

Outcome	Reported Estimate	Optimal Bandwidth	Estimate	
			No Covariates	Covariates
All teachers – NBCT	0.022 (0.007)***	0.17	0.007 (0.008)	0.013 (0.006)**
All teachers – NBPTS applicant	0.007 (0.002)***	0.36	0.008 (0.002)***	0.007 (0.002)***
All teachers – Experience	-0.133 (0.216)	0.28	-0.045 (0.227)	0.186 (0.130)
All teachers – Advanced degree	-0.005 (0.009)	0.44	-0.006 (0.008)	-0.001 (0.005)
All teachers – Turnover	-0.004 (0.005)	0.28	-0.005 (0.006)	-0.007 (0.006)
New Hires – NBCT	0.011 (0.006)**	0.70	0.010 (0.005)**	0.011 (0.005)**
New Hires – Experience	0.165 (0.285)	0.33	0.189 (0.291)	0.286 (0.279)
New Hires – Advanced degree	0.033 (0.014)**	0.70	0.027 (0.012)**	0.017 (0.011)
NBCTs – Turnover	-0.013 (0.012)	0.44	-0.014 (0.011)	-0.022 (0.011)**
All students – Math test	-0.005 (0.020)	0.36	-0.004 (0.020)	-0.014 (0.018)
All students – Reading test	-0.022 (0.016)	0.36	-0.022 (0.016)	-0.009 (0.013)

Notes: Estimated regression discontinuity effects with optimal bandwidth calculated for triangular kernel using the cross-validation procedure described by Lee and Lemieux (2010). Standard errors clustered by school in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.