Do School Report Cards Produce Accountability Through the Ballot Box?

Vladimir Kogan
Department of Political Science
Ohio State University
kogan.18@osu.edu

Stéphane Lavertu
John Glenn College of Public Affairs
Ohio State University
lavertu.1@osu.edu

Zachary Peskowitz
Department of Political Science
Emory University
zachary.f.peskowitz@emory.edu

January 8, 2016

Abstract

Public education has been transformed by the widespread adoption of accountability systems that involve the dissemination of school district performance information. Using data from Ohio, we examine if elections serve as one channel through which these accountability systems might lead to improvements in educational quality. We find little evidence that poor performance on widely disseminated state and federal indicators has an impact on school board turnover, the vote share of sitting school board members, or superintendent tenure, suggesting that the dissemination of district performance information puts little (if any) electoral pressure on elected officials to improve student achievement.

Keywords: performance measurement, school board elections, school district accountability, No Child Left Behind, Every Student Succeeds Act

1 INTRODUCTION

Public education has been transformed over the past two decades by the widespread adoption of accountability systems that involve the dissemination of school performance information. Elections are one channel through which such reforms might lead to improvements in educational quality, but the extensive scholarship on education accountability has largely ignored local politics as a potential mechanism.[1] This oversight is surprising because, despite recent centralization trends (Henig 2013), many key decisions about education policy and the day-to-day administration of public schools remain in the hands of local school districts whose boards serve at the pleasure of voters. If voters condition their support for school board members on observed performance and service quality, as is true in many other local electoral contexts (Arnold and Carnes 2012, Oliver and Ha 2007, Miller 2013), reforms that make district performance information accessible and salient could increase political pressure on these policymakers to improve student achievement.

This study examines whether the public dissemination of district performance information has had such an impact on local school board elections held from 2003 to 2012 across 613 Ohio school districts. The analysis employs panel data and regression discontinuity analysis to identify the causal impact of the federal "adequate yearly progress" (AYP) designation ("met" or "not met"), as well as five (and by the end of the sample period, six) state performance designations on school board turnover, the electoral success of incumbent school board members, and superintendent turnover. Ohio school districts provide a good opportunity to detect these effects because performance designations are released just prior to the November school board elections, the state includes a large number of school

[1]Research indicates that the No Child Left Behind (NCLB) Act of 2001 had some impact on test scores (Ahn and Vigdor 2013, Dee and Jacob 2011, Ladd and Lauren 2010), but we know little about the mechanisms responsible for this impact.

districts with performance designations that have varied over time, and the performance measurement schemes stayed in place for an entire decade (2003-2012). Indeed, research indicates that the federal AYP performance designations have had a significant impact on November school tax referenda in these very districts over the same time period (Kogan, Lavertu and Peskowitz 2015). Drawing on strong evidence that performance information disseminated via school "report cards" directly shapes voter perceptions about the quality of local schools (e.g., Chingos, Henderson and West 2012, Jacobsen, Saultz and Snyder 2013), we hypothesize that the likelihood of turnover increases among board members and superintendents when performance ratings signal low student achievement.

Yet the analysis reveals little evidence that publicized measures of school district performance have had an impact on the likelihood of turnover on school boards, the electoral success of sitting school board members, or the likelihood of superintendent turnover. Due to a lack of statistical power in some models, our findings are more tentative regarding the impact of state designations than they are for the federal AYP indicator. The evidence is strong, however, that AYP had a negligible impact on both board and superintendent turnover. These findings speak directly to the likely impact of the recently adopted Every Student Succeeds Act, which maintains existing requirements for the public reporting of academic performance information but eliminates almost all of the federal sanctions tied to poor performance. Our results are consistent with the findings of Hanushek and Raymond (2005), who suggest that performance information by itself produces little incentive to improve student achievement.

2 ACCOUNTABILITY SYSTEMS AND SCHOOL GOVERNANCE

2.1 School "Report Cards" in Ohio

There are approximately 14,000 school boards governing local school districts across the United States, and the vast majority of school board members are elected. These boards establish policies that govern the schools in their districts. However, the significant expansion of state and federal accountability systems in the 1990s and early 2000s, which imposed regular testing and performance reporting requirements on school districts, have come to play an increasingly important role in education governance. Most notably, the 2002 reauthorization of the federal Elementary and Secondary Education Act, known as the No Child Left Behind Act (NCLB), required all states to produce and disseminate yearly district and school "report cards" that identified, among other things, whether or not schools and districts made "adequate yearly progress" (AYP) toward attaining the federal policy goal of 100 percent student proficiency in mathematics and reading by 2014.

Like all states, Ohio report cards published from 2003 to 2012 labeled schools and districts as having "met" or "not met" federal AYP requirements. The binary AYP designation was important to administrators and educators because schools and districts failing AYP for two consecutive years entered "improvement status," which entailed the accumulation of increasingly invasive interventions and sanctions for every additional year of AYP failure. There is also evidence that voters used this information in local elections, as district AYP failure was associated with a lower probability of school tax referendum passage in Ohio—particularly in November elections that immediately followed the August release of school and district report cards (Kogan, Lavertu and Peskowitz 2015).

Like many states, Ohio chose to modify its state performance measurement system in

order to incorporate NCLB requirements, as opposed to administering parallel state and federal accountability systems. In addition to AYP, Ohio employed two (eventually three) other metrics in order to determine a district's overall state performance designation. Two performance measures used for the entire decade include an index aggregating student proficiency levels on state tests (in math, reading, writing, science, and citizenship) and a measure that captured the percent of state quality indicators met (focused on student proficiency, graduation, and attendance rates). Unlike the AYP indicator, which required that each and every student demographic subgroup within a school or district reach proficiency targets, these Ohio-specific metrics were based primarily on aggregate student proficiency rates and did not become more stringent over time. Finally, in 2008 Ohio introduced a measure of school and district "value added," based on relative, year-to-year student improvements on mathematics and reading exams in grades three through eight, as a fourth indicator of district quality.

These measures complemented one another in terms of how they accounted for various educational outputs, such as test participation requirements, attendance and graduation rates, and overall and subgroup performance on standardized tests. Depending on how schools and districts scored across the federal AYP and three state metrics (the performance index, percent of indicators met, and value added), they received one of the following designations, ordered from best to worst: "Excellent with Distinction," "Excellent," "Effective," "Continuous Improvement," "Academic Watch," and "Academic Emergency."[2] These overall performance designations were released every August, just prior to November school board elections. They received a great deal of attention in local media outlets and voters could also look them up on the Ohio Department of Education website. Addition-

[2]Until 2008, "Excellent" was the highest designation awarded. "Excellent with Distinction" was added in conjunction with the value-added measure.

ally, federal law required districts to notify parents if their children attended a school or district that failed to meet the federal AYP standard for two consecutive years or more.

2.2 School Board Governance in Ohio

The Ohio school boards we examine have the power to affect student learning via a number of mechanisms, such as updating district strategic plans, shifting the allocation of resources, modifying the procedures for hiring and evaluating the performance of teachers and administrators, negotiating with unions on the specifics of teacher contracts, and so on. Additionally, although boards in Ohio provide oversight and set general district policy, they may have a significant impact on day-to-day operations simply through their power to appoint, dismiss, or otherwise influence district superintendents and treasurers.

During our period of study, Ohio had 613 K-12 public school districts, with the median district in Ohio overseeing the operation of six schools serving 1,740 students. Elections are non-partisan and held every two years in November of odd years, with board majorities (4 seats in large urban districts and 3 seats in all other districts) up for grabs every four years.[3] The elections are at-large and voters can select as many candidates as there are seats available.[4] All adult U.S. citizens registered to vote within the boundaries of a school district may participate in school board elections.

[3]There are no term limits for school boards.

[4]Under Ohio law, candidates appear on the district ballot if they file a petition with 25-300 valid signatures, depending on the population within a district.

3 Causal Pathways

3.1 Electoral Pressure

Regular publication of school report cards should be expected to increase the electoral pressure on school board members to improve student achievement. First, by making performance information easily accessible, school report cards can make student achievement more salient in the minds of voters, increasing the weight put on this consideration when choosing between competing candidates. Second, lower information costs and higher salience may also have a mobilizing effect, increasing interest and participation among a broader, more representative segment of the electorate (Anzia 2011). Finally, negative performance information can make incumbent school board members appear vulnerable, encouraging the entry of stronger challengers (Jacobson and Kernell 1981, Levitt and Wolfram 1997, Stone et al. 2010) to offset the significant incumbency advantage in local nonpartisan contests (Krebs 1998, Shaffner, Streb and Wright 2001). Potentially operating through all of these channels, we expect performance information included on report cards to influence local electoral outcomes, with negative performance indicators reducing support for incumbent board members and increasing turnover.

To our knowledge, only two published studies have examined the effect of school performance measures on school board elections: Berry and Howell (2007) and Holbein (Forthcoming). Our study differs significantly from both. Berry and Howell study the impact of continuous, aggregate student test scores on school board elections in South Carolina, rather than the more widely publicized and accessible discrete federal and state performance designations considered here. Holbein uses a regression discontinuity design to examine the impact of the federal AYP metric on school board elections in North Carolina. Unlike our study, which focuses on turnover and electoral success among sitting school

board members in Ohio, Holdbein's outcome of interest is the margin of victory for all winning candidates, irrespective of their incumbency status. Because the incentives elected officials face depend on whether voters can accurately attribute credit and blame to the appropriate office holders (see, e.g., Healy and Malhotra 2013), this distinction is critically important. Both studies focus primarily on school-level performance, whereas we examine the effect of district-level designations because the policymakers we focus on—elected school board members and appointed superintendents—serve in district-wide offices. Finally, neither study includes an analysis of turnover among district superintendents, which one might expect if electoral incentives produce greater accountability pressures.

3.2 Report Cards and Superintendent Turnover

The dissemination of school performance information might affect the tenure of top school administrators for a number of reasons. First, voters might hold appointed district superintendents—rather than elected school board members—responsible for student achievement, in the same way that they appear to attribute credit and blame for economic performance to governors but not legislators (Atkeson and Partin 1995). This would create political pressure for board members to replace superintendents in low-performing districts. Second, school board incumbents who face electoral heat for poor performance may respond by removing the superintendent in an attempt to shift the blame.

4 EMPIRICAL STRATEGY

We hypothesize that, if student performance information helped voters hold school board members accountable for district educational outcomes, then indicators of poor district performance should be associated with (1) school board turnover, caused either by incumbent school board members losing their seats or retiring strategically in anticipation of tougher

contests, and (2) school board members dismissing superintendents in order to shift blame to the outgoing administrators. The empirical analysis examines how performance on the overall state rating system (up to six designations ranging from "Excellent with Distinctions" to "Academic Emergency") and the federal metric ("AYP met" or "AYP not met") affected the probability of turnover among school board members across all Ohio districts,[5] the share of votes won by sitting school board members in Ohio's largest metropolitan areas, and the probability of superintendent turnover across all of Ohio's districts.

We employed two strategies to identify these potential impacts. First, we estimated models with district fixed-effects to examine the impact of poor performance on board member turnover from 2003 through 2012 and superintendent turnover from 2004 through 2011. This strategy allows us to account for time-invariant unobservable differences between districts. Second, using our knowledge of how school and district performance designations were calculated, we employed regression discontinuity (RD) designs in which we compared turnover in districts that just failed to meet the performance standard for a particular designation to turnover in districts that just met that same performance standard. In our application, the academic designation of districts is determined by the continuous realization of student academic performance.[6] If districts have imperfect control over the determinants of their performance designations, an assumption that we test, then we can estimate the causal effect of these designations by examining how the outcomes of inter-

[5]Note that winning fewer votes conditional on running is only one mechanism that can result in higher turnover. Alternatively, school board members in low-performing districts might be less likely to run for re-election, resulting in more turnover (Berry and Howell 2007, Abernathy et al. 2015).

[6]Chakrabarti (2014) exploits similar variation to study the effects of AYP designations on future academic performance.

est change around the threshold (Lee 2008).[7] Although this RD approach yields results that may not be generalizable to all Ohio districts, it provides a useful complement to the fixed-effects models in that it relies on more plausibly exogenous variation in test-score performance. Conversely, the fixed-effects models provide insights about the generalizability of the RD results outside the neighborhood of the performance cutoff.

It is important to note that the RD and fixed-effects models do not estimate the same substantive quantities. The causal effects estimated with the RD approach are based on differences between districts immediately above and below the thresholds determining AYP status or a state designation. As such, the RD estimates are "local" treatment effects that capture a substantively small difference in actual student achievement. By contrast, the fixed-effects estimates capture "average" treatment effects, corresponding to the mean difference between districts receiving each performance designation.

Nevertheless, research indicates that voters should still be responsive to small changes in measured achievement if these changes affect the performance designations disseminated on state report cards. The official designations are far more salient and widely known than the underlying performance metrics that determine them, and research has found that small changes in performance indeed have an impact on perceptions of school quality. For example, using the RD approach, Chingos, Henderson and West (2012) find that very small changes in student test scores can produce substantial movement in citizen perceptions of school quality when such changes alter the grade awarded by the state to local schools. Other studies document similar effects, showing that arbitrary changes in performance standards can influence parental satisfaction with local schools (Jacobsen,

[7] As we elaborate below, our AYP RD application is actually a fuzzy RD design where the *probability* of achieving AYP status jumps discontinuously when the district crosses a test-score threshold (Imbens and Lemieux 2008, Lee and Lemieux 2010).

Saultz and Snyder 2013) and their willingness to make private financial contributions to support public education (Figlio and Kenny 2009)—even when these changes do not reflect any meaningful movement in actual student achievement.

4.1 Data

The analysis employs 2003-2012 Ohio district report card data and additional information available from the Ohio Department of Education. These data include variables that capture district performance designations, student demographics, information on district finances, and years in "program improvement" status due to AYP failure. Each of our models—predicting both school board and superintendent turnover—cover somewhat different samples of districts and years. Tables 1 and 2 report descriptive statistics for all of the variables used in each set of analyses. We describe these variables below.

Table 1: Descriptive Statistics for School Board Turnover

Variable	Mean	Std. Dev.	N
School Board Member Turnover	0.46	0.498	7320
District AYP Met	0.469	0.499	7320
Academic Emergency	0.008	0.089	7320
Academic Watch	0.035	0.185	7320
Continuous Improvement	0.227	0.419	7320
Effective	0.42	0.494	7320
Excellent	0.242	0.428	7320
Excellent w/ Distinction	0.067	0.249	7320
Years in Program Improvement	0.419	1.287	7320
Year 3 of Program Improvement	0.021	0.144	7320
> Year 3 of Program Improvement	0.048	0.215	7320
African-American (% of Enrollment)	5.767	14.816	7320
Hispanic (% of Enrollment)	1.564	3.259	7320
English Learners (% of Enrollment)	0.708	2.739	7320
Economically Disadvantaged (% of Enrollment)	29.662	18.462	7320
Total Expenditures (per Student)	8966.598	1975.717	7320

The ODE data files also included the district-level proficiency, graduation, attendance,

Table 2: Descriptive Statistics for Superintendent Turnover

Variable	Mean	Std. Dev.	N
Superintendet Turnover	0.161	0.368	4570
District AYP Met	0.465	0.499	4570
Academic Emergency	0.002	0.049	4570
Academic Watch	0.022	0.147	4570
Continuous Improvement	0.182	0.386	4570
Effective	0.441	0.497	4570
Excellent	0.277	0.447	4570
Excellent w/ Distinction	0.075	0.263	4570
Years in Program Improvement	0.474	1.322	4570
Year 3 of Program Improvement	0.026	0.161	4570
> Year 3 of Program Improvement	0.053	0.225	4570
African-American (% of Enrollment)	5.651	14.245	4570
Hispanic (% of Enrollment)	1.562	3.244	4570
English Learners (% of Enrollment)	0.773	2.931	4570
Economically Disadvantaged (% of Enrollment)	30.764	18.149	4570
Total Expenditures (per Student)	9123.877	1822.662	4570

and test participation rates necessary to calculate the state performance designations and federal AYP designations described above, as well as the names of district superintendents. The overall district performance designations were available for all years (2003-2012), but the data necessary to calculate some of the state and federal designations were missing for early years. Consequently, RD analyses for AYP are limited to 2004-2012. Additionally, the state "value-added" metric, which we use to conduct the RD analysis for state designations, was introduced in August 2008.

The performance measures that we study exhibit a great deal of within-district variation over the sample period. Of the 613 school districts that we observe for at least one year in the sample,[8] 512 undergo a change in AYP status while 26, 81, 411, 517, 471, and 236 districts undergo a change from the Academic Emergency, Academic Watch, Con-

[8]Some districts merged during our sample period, so there are now 611 districts in Ohio.

tinuous Improvement, Effective, Excellent, and Excellent with Distinction designations, respectively. Very few districts maintain the same performance designation over the course of the sample and we exploit this variation in the fixed-effects specifications.

We coded 2003-2012 school board membership using rosters compiled by the Ohio Secretary of State. To measure board turnover, we compare the roster in the year of each election—2003, 2005, 2007, 2009, and 2011—to membership lists for years immediately following the election, identifying individuals whose original term was up in the election year and whose name no longer appears in the latter publication.[9] We also obtained election records from 27 county boards of elections in order to code school board election results across Ohio's largest metropolitan areas (Cincinnati, Columbus, Cleveland-Akron, and Dayton). Finally, we obtained publicly available information on superintendents from the Ohio Department of Education and Ohio Treasurer's websites. As Tables 1 indicates, we observe school board turnover in approximately 46 percent of the 7,320 member-by-election-year observations in the sample. In contrast, as can be seen in Table 2, superintendents separate from their districts in approximately 16 percent of the 4,570 district-years. Note that the most reliable data on superintendent turnover were available from the 2003-04 through 2010-11 school years, so the results reported below are limited to those years.

Table 3 presents some preliminary descriptive statistics on the relationship between district performance and turnover. It reports the aggregate, cross-sectional school board and superintendent turnover rates across all district-years and for each federal and state designation category. With one exception, the table reveals no clear correlation between district performance and the tenure of elected officials or appointed administrators—a general pattern that holds in our panel regressions. The one exception is higher turnover

[9]Note that not all school board members have a term that ends in each election year, so we can observe turnover only for half of the members, unless someone resigns mid-term.

observed for superintendents representing districts receiving the lowest Academic Emergency designation. However, these districts differ substantially from all others in terms of the socioeconomic composition of their students. After accounting for these socioeconomic differences between districts in our multivariate models, we find no consistent relationship between state designations and superintendent turnover.

Table 3: Cross-Sectional Relationship Between State and Federal District Designations and Concurrent Turnover

	School Board Turnover	**Superintendent Turnover**
AYP Met	45.1%	15.4%
AYP Not Met	46.7%	16.8%
Pearson's r	-0.01	-0.02
Academic Emergency	42.4%	36.4%
Academic Watch	56.0%	17.9%
Continuous Improvement	48.9%	18.1%
Effective	45.4%	16.5%
Excellent	44.8%	14.5%
Excellent w/ Distinction	39.1%	14.0%
Pearson's $r^{\dagger}$	-0.05	-0.04

† The correlation is calculated by combining the state designations into a linear index.

4.2 Statistical Models

The empirical analysis employs both fixed-effects and regression discontinuity specifications. For all specifications, we examine turnover at time t using performance information disseminated in the same calendar year (t) and in the previous calendar year ($t-1$). For superintendent turnover, we also examine performance information from two years prior ($t-2$). As we explain below, we expect the largest effects for school board elections and turnover to occur in the same year (t), when recent performance is most salient in the minds of voters. For superintendent turnover, however, we remain agnostic about the precise timing of the effects.

The fixed-effects specifications exploit the panel structure of the data to remove unobserved district heterogeneity. The following equation presents our specification for the superintendent turnover models:[10]

$$\text{Turnover}_{it} = \beta_1 \text{AYP Met}_{it} + \boldsymbol{\beta_2}\textbf{State Designations}_{it} + \gamma\textbf{Controls}_{it} + \theta_i + \alpha_t + \epsilon_{it}$$

Our key quantities of interest are β_1 and $\boldsymbol{\beta_2}$, corresponding to the effects of meeting AYP and receiving one of the state performance designations (captured with a set of indicator variables), respectively. The district-specific effects, θ_i, capture time-invariant differences between districts. The fixed-effects specifications exploit only within-district variation, so it is less plausible that unobserved district characteristics would be falsely driving the association between performance and turnover. On the other hand, these specifications rely on a conditional independence assumption to identify the causal effect of performance

[10]For school boards, the turnover outcome variable has an additional subscript for each board member, although all of the variables on the right-hand side of the equation continue to be measured at the district level, and the errors are also clustered at this level.

information on turnover and incumbent vote share. If there are time-variant unobservables that are correlated with the performance information and our outcomes of interest, then our coefficient estimates will be biased. We include a set of covariates, $\mathbf{Controls}_{it}$, that account for time-varying district characteristics listed in Tables 1 and 2 as well as year fixed-effects α_t, and we report heteroskedasticity robust standard errors that are clustered at the district level.

To complement the fixed-effects results, we also report coefficient estimates from regression discontinuity models. In the case of a "sharp" RD design, all districts with realizations of a running variable above an arbitrary performance threshold receive treatment (one of two adjacent performance designations) and districts below the threshold are assigned to the control condition (the other performance designation). As we explain below, our analysis of the six state designations employs a sharp RD design. However, in the case of the federal AYP designations, treatment status is not a deterministic function of the realization of the running variable because we do not use all of the information that determines a district's performance designation. The running variable we created leads us to calculate the AYP designation with error. Consequently, we estimate "fuzzy" RD models instead.

In the fuzzy RD setting, functions of the realization of the running variable can be used as instruments for the performance designations and the causal effect can be estimated by the just-identified instrumental variables estimator or two-stage least squares in the overidentified case. In the results we present below, we follow the advice of Gelman and Imbens (2014) and instrument for treatment status with the linear and quadratic form of the running variable, interacted with an indicator for whether the running variable takes on a value greater than zero. Equipped with predicted treatment status, $\hat{\tau}_i$, our second

stage regression is the following:

$$\begin{aligned}\text{Turnover}_{it} = \hat{\tau}_{it} + \beta_1 X_{it} + \beta_2 X_{it}^2 + \beta_3 X_{it} * 1_{X_{it} \geq 0} \\ + \beta_4 X_{it}^2 * 1_{X_{it} \geq 0} + \gamma \mathbf{Controls}_{it} + \epsilon_{it}\end{aligned}$$

As in the first stage, this specification allows the probability of turnover to vary as a quadratic polynomial of the running variable and the functional form of the polynomial is permitted to vary to the left and right of the cutoff. In most specifications we report, we also included additional covariates, $\mathbf{Controls}_{it}$. In all of our RD specifications we report heteroskedasticity robust standard errors that are clustered at the district level for the AYP models and clustered at the district-year for the models focused on the state designations.[11]

4.3 Constructing Running Variables for the RD Analysis

District AYP designations are based on a set of necessary conditions. The failure to meet any one AYP requirement results in a designation of "AYP not met." Consequently, one can create a one-dimensional running variable capturing district performance relative to AYP targets by focusing on the AYP requirement with respect to which a district performs most poorly. Scholars have employed this "minimum distance" technique specifically as it relates to NCLB's AYP metric (see Chakrabarti 2014).

The factors used to determine each district's AYP designation include reading and math

[11]This is particularly important in our school board candidate vote-share specifications reported in the Supplemental Appendix, where the outcome of interest is the share of votes won by incumbents. Because multiple incumbents often compete simultaneously and potentially against each other in the same election we code each incumbent's share of all votes cast.

proficiency rates, test participation rates, attendance rates, graduation rates, and, eventually, achievement growth. We constructed variables that capture the difference between the realized district performance on each of these dimensions and the performance required for meeting AYP (e.g., the district participation rate minus the 95 percent participation rate required for AYP) for each of 10 student subgroups[12] and in each subject area (when applicable) for each district-year.[13] Thus, the running variable captures the lowest value of realized performance for each district-year. This minimum value is negative for districts that failed at least one AYP requirement and positive for districts that satisfied all of them.

Because almost all districts that failed to meet AYP usually did so due to the proficiency requirements, we provide a concrete example of the calculation using these criteria. Suppose that in one year, 65 percent of students were required to be proficient in both math and reading for a district to meet AYP. Note that this goal applied not only to the overall student body but also to each of the student demographic subgroups. Thus, we first calculated the proficiency rate for each of these subgroups in both math and reading.[14] Then, we calculated the difference between each district subgroup's proficiency rate and the performance requirement (i.e., the subgroup proficiency rate in each subject minus 65 percent) in math and in reading. Because failing to reach the threshold in a single subject by just one subgroup was sufficient for a district to fail to meet AYP, the running variable

[12]In addition to the overall student body, these include white, African-American, Latino, Asian, American Indian, and multiracial students; those who qualify for free and reduced lunch; students receiving special education services; and English learners.

[13]Federal regulations allowed districts to use either their one-year or two-year proficiency totals, and for each subgroup and subject, we identified the standard that put that subgroup on the best possible footing.

[14]As we note above, both the one-year and two-year averages are used.

we constructed used the calculated difference—proficiency rate minus 65 percent—for the lowest-performing student subgroup-subject combination. For example, if a school district's lowest achievement is a proficiency rate of 40 percent in math among its "economically disadvantaged" subgroup, the value of the district's running variable in that year would be -25 (40 percent minus 65 percent).

We used an analogous procedure for calculating the running variables for the state performance designations. The primary difference is that the state designations were based on aggregate performance — rather than the achievement of student demographic subgroups — and initially incorporated only two, and later three, separate requirements.[15]

A threat to identification in the context of the regression discontinuity design is the ability of agents to manipulate the realization of the running variable. If school district administrators or board members have the ability to precisely manipulate their district's performance rating, then treatment status may be confounded with unobservable district characteristics that could be driving the relationship between our outcomes and district performance designations. To check for manipulation of the running variables, we used the density test developed by McCrary (2008). Our preliminary analysis, which employed running variables that perfectly replicated the assignment of federal and state performance designations, showed significant discontinuities in densities at the performance cutoff for each pair of adjacent designations—that is, where the running variables take on a value of zero. That districts near a performance threshold are significantly more likely to receive the better of two adjacent designations violates the assumptions underlying the RD design (McCrary 2008).[16]

[15]In addition, it was sufficient to meet just one of two of these requirements to receive each state designation, so we used the highest of the two values for the running variable, rather than the lowest as was the case for the AYP running variable.

[16]One explanation for the manipulation is that districts strategically tried to game the

We addressed this problem in two ways. First, we removed several problematic underlying measures[17] from the AYP calculation and used a "fuzzy" RD design to deal with the imperfect calculation of the designation based on the subset of measures that show no sign of manipulation. More precisely, we created a running variable based solely on the subgroup proficiency rate requirements. Unlike the other indicators used in the AYP calculation, it was harder for districts to *precisely* manipulate student proficiency rates.[18] As a result, relying only on the participation goals eliminated discontinuity of the running variable at the threshold[19] and the new variable remained a very good predictor of a district's ultimate AYP classification. As illustrated in Figure 1, the probability that a district meets AYP jumps by about 32 percentage points when a district passes the running variable threshold of 0. That is the variable we use in our analysis.

Second, for the state designations, we created a running variable that employs only the metric showing no signs of manipulation: district academic value added. This measure captures the year-to-year achievement growth of individual students. Recall that, until 2008, the state assigned each district a performance rating based on two indexes of aggregate student achievement (a "performance index" and the percentage of "quality indicators" met).

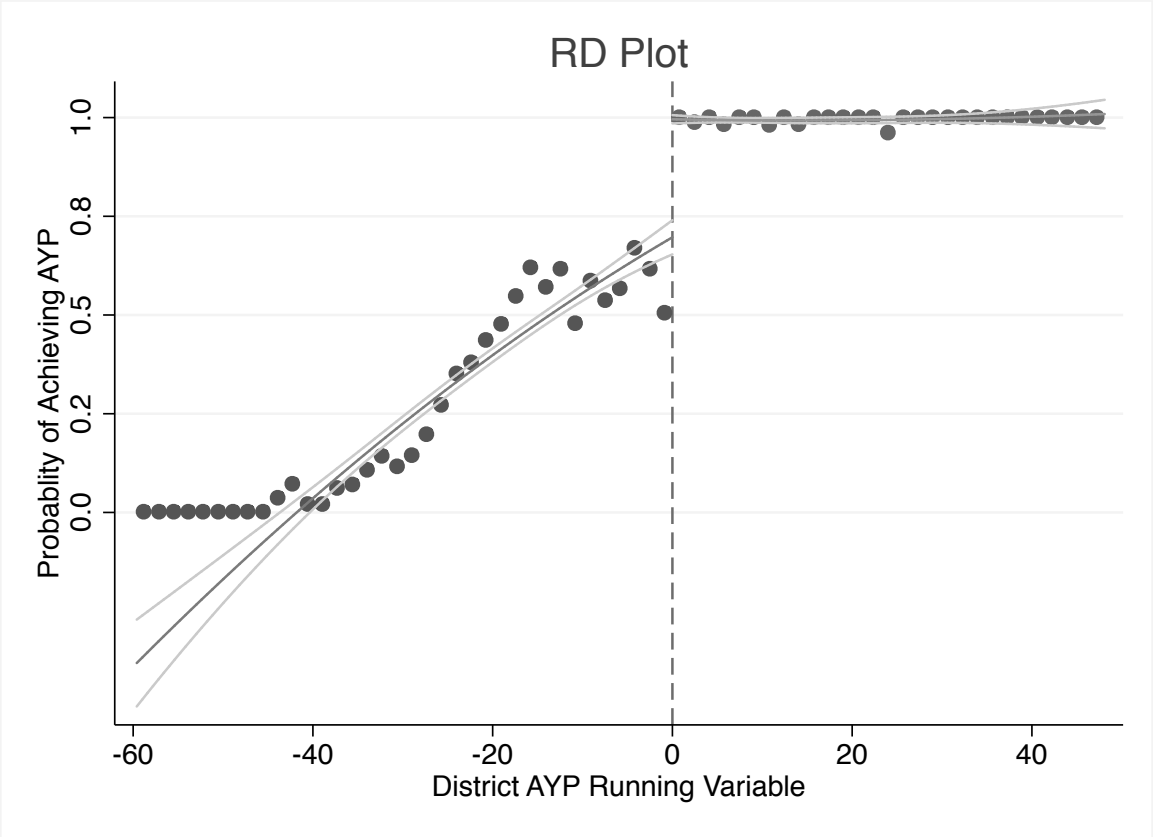
test scores—by, for example, disenrolling and then re-enrolling low-performing students—to improve their ranking.

[17]We excluded the attendance, graduation, and participation rates, in addition to the "safe harbor" provision that exempted districts posting significant improvements in their proficiency rates from year to year.

[18]While districts could still affect measured student proficiency in less precise and predictable ways, the types of strategies available to do so could not be applied only at the threshold.

[19]The Supplemental Appendix reports the McCrary tests for the modified running variables.

Figure 1: Running Variable and Probability of Meeting AYP



Note: The figure plots the probability of a district meeting the federal Adequate Yearly Performance standard as a function of our running variable. The dots are binned averages and the dark line is a best fit line from a regression with a quadratic polynomial.

However, beginning in 2008, the rating assigned based on these indexes could be adjusted up or down by one designation depending on growth in students' achievement in math and reading. When a district's students consistently posted more than one year's worth of academic growth, its state designation was increased to the next highest rating.[20] For

[20]The discussion above is somewhat of a simplification, as not all districts were eligible for this adjustment in all years.

example, in 2010, a district initially assigned the "Continuous Improvement" designation would be bumped up to "Effective" if it had two consecutive years of students recording more than one year's worth of growth. Similarly, districts in which students recorded less than one year's worth of achievement growth for three consecutive years had their designations reduced by one unit. For example, a district labeled "Continuous Improvement" based on the pre-2008 procedure would have its designation reduced to "Academic Watch."[21]

For each year in the dataset, we calculated two quantities. The first was the difference between a district's value-added index score and the minimum threshold to be classified as having achieved more than one year of growth (i.e., a district's value-added index score minus 1 or its index score minus 2, depending on the year). The second was the difference between a district's value-added index score and the maximum score that would still result in it being classified as having achieved less than one year of growth (a district's value-added index score minus -1 or its index score minus -2, depending on the year). In the first case, a value above zero indicates that a district's students had more than one year of achievement growth. In the second case, a value above zero indicates that a district avoided being labeled as having less than one year of achievement growth.[22] Because multiple consecutive years of above- or below- expected growth were necessary to see either adjustment in the designations, we constructed the running variables using the value-added index scores from the year that was most consequential for affecting each district's designation. For the upward adjustments, the running variable is thus based on the year

[21]Note that the number of years of high or low growth needed for these adjustments changed over the course of the period we examine.

[22]Note that is also a residual category—which included districts with index scores in the interval between these positive and negative thresholds—for districts that were classified posting neither more nor less than a year of growth.

during the relevant period of time in which its students made the least amount of growth. For the downward adjustments, the running variable captures each district's highest year of student growth. More specifically, the running variable for upward adjustments captures the smallest calculated difference (gain index score minus 1 or 2, depending on the year) over a period of several years, whereas the running variable for downward adjustments captures the largest calculated difference (gain index score minus -1 or -2, depending on the year) over the relevant period. A district with a positive value of either running variable received a higher designation than it otherwise would have in the absence of the adjustments.

As an example, suppose that a district in 2010 would have been assigned the "Continuous Improvement" designation, based on the pre-2008 procedure. To move up to "Effective," the district would need two years of exceeding the value-added goal. Its lowest score in those two years is the value of the running variable for this shift, because if the score in that year exceeded the threshold, the district necessarily would have exceeded the goal in the year when its students posted higher achievement growth. To move down to "Academic Watch," the district would need three years of low growth, so, for the same reason, we use its highest scoring year as the running variable for this shift. In both cases, moving from below zero to above zero on the running variable means the district moves from the lower to the higher of the two designations, from "Continuous Improvement" to "Effective" and from "Academic Watch" to "Continuous Improvement," respectively.

To be clear, a district could receive only one of these value-added adjustments in determining its state designation in any given year, because districts could not simultaneously post both more *and* less than one year of growth in any single year. Many districts received neither adjustment, because their value-added index scores were not consistently high or low across multiple consecutive years. The data we analyze include two observations for each district in each year—corresponding to the running variables for adjustments

up or down—and we cluster the standard errors in the state RD models by district-year to account for this fact.

Because we have access to all of the necessary value-added calculations, we use the "sharp" RD design for our analysis of the state designations.[23] This portion of our analysis is limited to 2008-2012, the period for which the adjustments applied. In the results presented below, we pool all of the state designations for the RD analyses to maximize our statistical power. Doing so assumes that a one-unit increase in a state designation produces the same effect on turnover regardless of the baseline designation. In other words, the pooled analysis assumes that improving from "Academic Watch" to "Continuous Improvement" should produce the same effects as improving from "Excellent" to "Excellent with Distinction." However, as we show in Tables A.3 through A.5 in the Supplemental Appendix, relaxing this assumption and allowing the effect to vary across designations does not substantively change the results. The Supplemental Appendix also reports the results of density and covariate balance tests for these modified running variables.[24]

5 EMPIRICAL RESULTS

The analysis examines the impact of performance designations on school board turnover and superintendent turnover using the fixed-effects and regression-discontinuity designs described above. Given evidence that voters update their beliefs about district quality almost immediately after new indicators are released (e.g., Chingos, Henderson and West

[23]The analysis excludes six district-year observations for which we were unable to reconstruct the running variable.

[24]All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at `http://www3.interscience.wiley.com/cgibin/jhome/34787`.

2012, Kogan, Lavertu and Peskowitz 2015) and that the effect of such information decays quickly over time (Healy and Lenz 2014), we expect district performance information released in August to be most salient in that year's November election. Thus, "concurrent" models—those examining the impact of designations released in the same year that turnover is observed—are most likely to identify an effect of performance information on school board turnover. However, we also report the results of models with a one-year lag of these designations.

In contrast, there is likely to be some lag between the release of annual performance designations and their effects on superintendent turnover because most superintendent contracts begin at the start of the fiscal year (July 1 of each year), before the performance designations are made public in August. Superintendents may also have multiple years left on those contracts.[25] Therefore, for the analysis of superintendent turnover, we present models that include concurrent designations as well as one- and two-year lags.

The variable AYP Met indicates that a district met the AYP target. The variables for the state designations indicate which of the listed designations a district received relative to the omitted category of Academic Watch. The indicator variables for the state designations are listed so that they begin with the worst designation, Academic Emergency—the designation just below Academic Watch—through to the highest Excellent w/ Distinction designation.[26] Better State Designation indicates whether or not the "value added" metric shifted the district's preliminary state designation up by one unit.

[25]Of course, these contracts can be terminated early.

[26]The results are identical if we instead use Academic Emergency as the omitted category. We prefer Academic Watch, the second-lowest designation, because more observations fall into this category, allowing us to produce more precise estimates.

5.1 School Board Turnover

Table 4 presents fixed-effect estimates of the impact of performance information on school board member turnover. The dependent variable indicates whether (1) or not (0) a sitting school board member whose term is expiring that year no longer appears on the school board roster in January of the subsequent year. Across all specifications, the magnitudes of the AYP coefficient are substantively small and statistically indistinguishable from zero. The magnitudes of the coefficients for the state performance designations tend to be larger, but they are never significant at the 5 percent level. In addition, many of the coefficients for the state designations are positive, indicating that moving from Academic Watch to better designations *increases* the probability of school board member turnover, the opposite of the theoretically-predicted relationship. Most strikingly, receiving the lowest Academic Emergency designation is consistently associated with less turnover compared to the higher Academic Watch designation. As we discuss in the conclusion, this may reflect a desire on the part of voters to avoid further academic disruption due to turnover.

We now turn to the relationship between concurrent performance information and school board member turnover in the regression discontinuity framework. Figure 2(a) presents the relationship between the federal Adequate Yearly Progress running variable and school board member turnover. The figure reveals that the probability of school board member turnover is almost constant across the realization of the AYP running variable and that there is no evidence of a statistically significant decrease in school board member turnover for districts that just cross the probabilistic treatment status threshold.[27] In fact, the probability of school board member turnover actually increases very slightly for districts that cross the threshold, but this effect is statistically indistinguishable from zero.

[27]As discussed above, the probability of treatment status, meeting the AYP standard, jumps at the 0 threshold in our fuzzy regression discontinuity design.

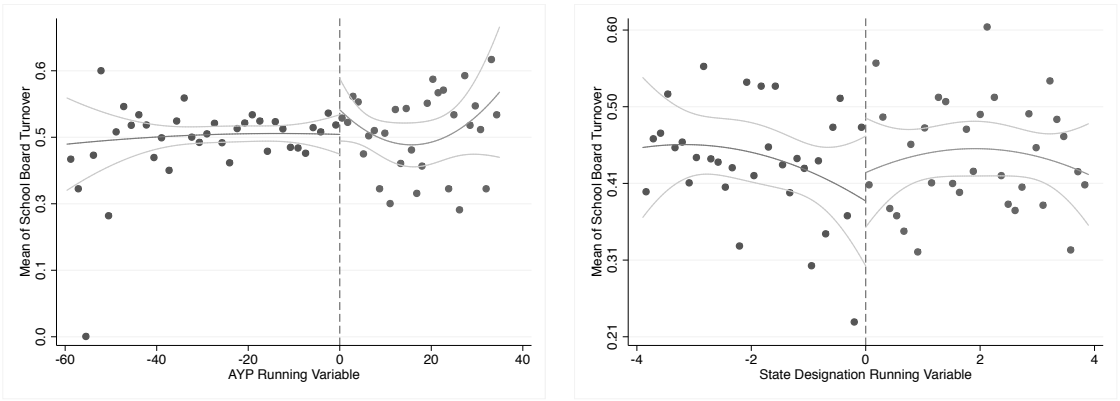
Table 4: District Performance Information on School Board Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.00496	0.00929	-0.0113	0.0224
	(0.0156)	(0.0191)	(0.0192)	(0.0246)
Academic Emergency	-0.120^{+}	-0.164	-0.0462	-0.0782
	(0.0728)	(0.122)	(0.190)	(0.343)
Continuous Improvement	-0.0278	0.0528	0.0631	0.155^{+}
	(0.0464)	(0.0561)	(0.0664)	(0.0853)
Effective	-0.0862^{+}	0.0334	0.0188	0.130
	(0.0509)	(0.0633)	(0.0703)	(0.0927)
Excellent	-0.0699	0.0461	0.0517	0.175^{+}
	(0.0552)	(0.0687)	(0.0750)	(0.0989)
Excellent w/ Distinction	-0.0990	-0.0344	0.0475	0.160
	(0.0644)	(0.0777)	(0.0836)	(0.109)
Observations	7320	7320	5910	5910
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Similarly, Figure 2(b) presents the results for the adjustments in Ohio state performance designations attributable to the value-added metric. The estimated effect of having a better state designation on school board member turnover is negligible at the cutoff.

These graphical relationships are confirmed in Tables 5 and 6, which present the RD regression results using both the full sample and a local sample, which includes only obser-



(a) Federal Adequate Yearly Progress running variable and incumbent school board member turnover

(b) "Better state designation" running variable and incumbent school board member turnover

Figure 2: The figure displays the relationship between the performance information running variables and the probability of turnover among incumbent school board members. The circles indicate the local average school board turnover; the dark lines represent the quadratic best fit estimated separately to the left and the right of the threshold; and the 95 percent confidence interval is demarcated using light lines.

vations within the bandwidth calculated using the Calonico, Cattaneo and Titiunik (2014) algorithm.[28] Once again, we find no evidence that lower performance ratings increase turnover among school board members. Although the coefficient on *Better State Designation* is positive and marginally significant in some specifications, this indicates greater turnover among better performing districts.

To examine the role of elections in bringing about school board turnover after accounting for strategic retirement decisions, we also estimated the impact of federal and state performance designations on election outcomes among school districts in Ohio's largest metro areas. These results are reported in Tables A.6 through A.8 in the Supplemental

[28]For the RDD estimates of the twice-lagged superintendent turnover specifications, we do not have enough observations to compute the optimal bandwidth. For these, we restricted the analysis to (-5,5) for the local sample.

Table 5: District AYP Passage on School Board Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.166	-0.0601	-0.0820	0.214
	(0.127)	(0.142)	(0.130)	(0.180)
Observations	5910	1336	5910	1429
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2004-2012. Heteroskedasticity robust standard errors clustered at district level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Better State Designation on School Board Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
Better State Desig.	0.0572^{+}	0.0342	-0.00818	-0.00608
	(0.0303)	(0.0562)	(0.0323)	(0.0577)
Observations	5524	3140	3823	2531
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix. These models provide no evidence of accountability pressures produced through elections.

5.2 Superintendent Turnover

To explore the impact of performance information on administrative turnover, we estimated models in which the dependent variable indicates whether the superintendent in place for the school year at time t is different than the superintendent in place during the previous school year. As noted above, models featuring concurrent performance designations (i.e., performance information released in year t) are unlikely to reveal evidence of blame-shifting because the district report cards are released after the current fiscal year has begun. Additionally, the search for a replacement superintendent can take time, so one probably should not expect a concurrent effect even if districts could respond immediately to report card information. Instead, we expect to find an effect in the models with lagged performance designations. In Tables 7 to 9, we present the fixed-effect and regression discontinuity results for superintendent turnover. Once again, we find no evidence that worse performance increases the probability that the district superintendent changes, echoing the findings from the school board turnover and vote-share models.

Table 7: Performance Information on Superintendent Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	0.0149	0.0156	0.0134	0.0135	0.00190	-0.0175
	(0.0137)	(0.0175)	(0.0140)	(0.0176)	(0.0154)	(0.0206)
Academic Emergency	0.147	0.221	-0.0378	-0.0590	0.0388	0.0765
	(0.161)	(0.195)	(0.0962)	(0.129)	(0.109)	(0.123)
Continuous Improvement	0.0353	0.0141	0.0110	-0.0241	0.0924^{+}	0.0452
	(0.0535)	(0.0671)	(0.0473)	(0.0621)	(0.0472)	(0.0723)
Effective	0.0274	0.0196	0.0286	0.00766	0.0743	0.0291
	(0.0595)	(0.0747)	(0.0516)	(0.0678)	(0.0525)	(0.0802)
Excellent	0.0387	0.0601	0.0249	0.0193	0.0914	0.0508
	(0.0634)	(0.0791)	(0.0552)	(0.0719)	(0.0575)	(0.0857)
Excellent w/ Distinction	0.0376	0.0344	0.0339	0.0137	0.114^{+}	0.0831
	(0.0678)	(0.0835)	(0.0624)	(0.0791)	(0.0686)	(0.0999)
Observations	4570	4570	4570	4570	3997	3997
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: District AYP Passage on Superintendent Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	-0.0513	-0.0294	-0.140	-0.0851	-0.0993	-0.0691
	(0.0672)	(0.0691)	(0.0883)	(0.103)	(0.0773)	(0.114)
Observations	3999	1633	3999	1633	3418	1271
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. Heteroskedasticity robust standard errors clustered at district level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Better State Designation on Superintendent Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
Better State Desig.	0.0178	0.0249	0.0118	0.0472	0.0263	0.0443
	(0.0245)	(0.0481)	(0.0274)	(0.0614)	(0.0339)	(0.0586)
Observations	3727	1947	2618	1132	1539	1000
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3 Sensitivity Analysis

We conducted numerous sensitivity analyses and report some of them in the Supplemental Appendix. As we note above, we examined the impact of performance information on the vote shares of sitting school board members in Ohio's largest metropolitan areas (Tables A.6-A.8). But we also examined, for example, the probability of board and superintendent turnover using non-linear, conditional logit models (Tables A.18 and A.19); estimated RD models separately for each state designation (Tables A.3-A.4); re-estimated all fixed-effects models after including additional controls capturing students' proficiency rates in mathematics and reading (Tables A.9-A.11); re-estimated all models so that they are limited to observations for which the AYP and state designations offer consistent signals of school district quality (Tables A.20-A.23); and re-estimated all superintendent turnover models focusing only on long-serving superintendents (Tables A.15-A.17). These analyses yield qualitatively similar results.[29]

One potential limitation of our analysis is that it focuses on district-level performance designations, whereas voters might be more concerned with the performance designations associated with their neighborhood schools—particularly if they have school-age children. We explored this possibility in a number of ways. First, we re-estimated our models so that they account for the percent of schools in the district receiving positive designations (Table A.1) and the percent of students in a district who are enrolled in schools receiving positive designations (Table A.2). Second, we re-estimated all models to look for differences in effects between large and small districts, which we categorized based on whether they

[29]Several specifications yield results suggesting that receiving the Academic Emergency designation is associated with lower incumbent vote shares and more superintendent turnover, but there are very few districts receiving this designation and the results for the other designations are generally not consistent with our accountability hypotheses.

had higher or lower enrollment than the median district in 2003 (Tables A.12-A.14). The logic of this second analysis is that the performance designations for these small districts are more likely to capture the performance of schools of concern to voters. Both of these analyses yield results that are qualitatively similar to those we present above.[30]

5.4 Summary of Results

Overall, we find minimal evidence that voters hold board members accountable for district performance on federal and state measures. There are no consistent results indicating that federal and state designations affect school board or superintendent turnover.[31] Indeed, many of our estimates for the school board turnover (and vote share) models have the opposite sign than one would expect if incumbents were punished by voters for poor performance. To the extent that any consistent effects emerge from the analysis, they are at odds with the expectation that indicators of low performance should produce greater turnover.

It should be acknowledged, however, that some of our estimates are imprecise, with standard errors that span very wide intervals. This is particularly a concern for the two lowest state designations, because there are very few districts moving in and out of these designations during the panel. In addition, some of our RD results are imprecisely esti-

[30]The one exception is that small districts receiving the Academic Emergency designation are associated with more board member turnover. However, there are very few small districts that receive this designation and we find no similar effects for the other state designations.

[31]The results are the same if we focus only on districts with long-serving superintendents, who can most credibly be blamed for performance outcomes. These results are presented in Tables A.15 through A.17 and A. 23 in the Supplemental Appendix.

mated. To address the problem of statistical power for the state designations, we have also re-estimated all fixed-effect models with a single additive index capturing each district's state performance rating. The index ranged from 1 for districts in "Academic Emergency" to 6 for those receiving the "Excellent with Distinction" designation. In these specifications, the coefficient on the state index remained statistically insignificant but was much more precisely estimated than the individual state rating coefficients, allowing us to rule out substantively large effects.[32]

We also note that we have sufficient statistical power to rule out all but the smallest, substantively negligible effects for meeting the federal AYP benchmarks across all of the fixed-effects models and can do the same for the state designations in some (but not all) of the RD specifications. Overall, we can consistently reject AYP effects on the order of one-twentieth of a standard deviation in school board and superintendent turnover, and much smaller effects in some specifications.

One additional concern is that, because AYP status is one of the measures used to determine each district's overall state designation, including these designations as controls in the AYP models might attenuate the estimated effect of meeting AYP because of post-treatment bias. To address this concern, we have also re-estimated all AYP models after excluding the state designations. In these specifications, the AYP coefficients remained very small, statistically indistinguishable from zero, and precisely estimated.

To summarize, we find no evidence that voters act on these state or federal performance designations nor that school boards respond to them when making district staffing decisions, and in many specifications we can rule out all but the smallest effects with a high degree of confidence. Thus, although our findings are more tentative for the state designations than for the federal AYP standard, the consistency of the findings across alter-

[32]These results are available from the authors.

native specifications, empirical strategies, and outcome measures all suggest that, at best, performance ratings produce modest political pressure to improve student achievement.

6 CONCLUSION

We hypothesized that voters would be less supportive of incumbent school board members and that the likelihood of turnover among board members and superintendents would be higher when performance metrics indicated low student achievement. Yet, using rigorous analytical strategies that minimize the possibility of omitted variable bias—and focusing on a state that in many respects enhances the likelihood of identifying an impact—the results indicate that publicized measures of school district performance have had little impact on the turnover of school board members or their top district administrators. Due to a lack of statistical power in some models, our findings are more tentative for the state designations than they are for the federal AYP indicator. But the evidence—particularly from the fixed-effects specifications—is strong that AYP had a negligible impact on board and superintendent turnover.

Research has found that test score-based performance designations correlate well with parents' own evaluations of their children's schools (Charbonneau and Van Ryzin 2012) and that parents update their evaluations to reflect new test score results (Chingos, Henderson and West 2012). It also indicates that the federal AYP metric had a significant impact on the behavior of Ohio voters. Specifically, Kogan, Lavertu and Peskowitz (2015) find that the probability of school tax referendum passage was approximately 10 percent lower in districts that received an "AYP not met" designation. In other words, there is evidence that Ohio voters responded to these performance metrics in the very same elections examined in this study. There is also anecdotal evidence that Ohio's performance designations were salient to Ohio citizens. Newspaper accounts, for example, regularly referenced the performance

designations when reporting on Ohio school districts.

Thus, indications are that citizens generally care about educational quality as measured by student achievement on standardized exams, and that voters in Ohio acted on this information when voting on ballot measures dealing with district funding. One plausible explanation for the null results of this study, therefore, is that voters were unable to link the performance of their districts to incumbent school board members. Ohio ballots provide no information about the incumbency status of candidates. Whereas Ohio voters could easily draw a connection between district performance and proposed tax referenda, they may have found it more difficult to link performance to school board members because it was far from obvious which candidates to hold accountable.[33] Another possibility is that, although test scores affect voters' perception of local school quality, they do not represent the only or even predominant consideration on Election Day. As Jacobsen, Snyder and Saultz (2014) show, many parents also place great value on their children receiving a well-rounded education, balancing high academic standards with exposure to the arts and other experiences. Thus, as states and the federal government have increasingly prioritized outcome-based accountability, voters may have come to look to school board members to balance accountability measures with other educational goals.

Finally, it is also possible that voters chose not to hold all incumbents equally accountable for low performance. Perhaps some sitting school board members could credibly claim the mantle of reform and build their reelection campaign around a platform of turning around low-performing districts. It is not possible to empirically test these possibilities with the available data, but they are inconsistent with a great deal of existing evidence

[33]The fact that multiple school board members govern a single district, and share their authority with appointed superintendents, creates additional hurdles for voters in terms of accurately attributing credit and blame for measured performance.

on retrospective voting. These studies show that voters regularly hold incumbents accountable for negative policy outcomes, including in domains over which these officials have little direct control (e.g., Miller 2013). They also indicate that voters struggle to accurately attribute credit for observed policy performance, frequently holding the wrong officials responsible for realized outcomes (e.g., Brown 2010).

Alternatively, some voters may have been hesitant to remove sitting board members, perhaps because they put greater weight on experience in the face of disappointing achievement or because they sought to avoid further academic disruption due to turnover. This explanation is inconsistent with our overall findings. While we do find some significant coefficients in a subset of models corresponding to less turnover when achievement is low, we do not find such effects consistently across all specifications. Nevertheless, this is a possibility that should be examined in future research.

Overall, this study indicates that despite the wide dissemination of simple and clear performance information, there is little evidence that electoral pressure served as a mechanism that motivated school board members to improve the quality of public education in Ohio. This is consistent with Hanushek and Raymond (2005), who found that sanctions are necessary for accountability systems to create incentives to improve student achievement. However, the negligible electoral impact in Ohio may not be generalizable to states with political institutions that help voters hold incumbents accountable for observed service quality, perhaps because their elections provide voters with more information about candidates.

References

Abernathy, Claire, Jason A. Grissom, Marc Meredith and James Sadler. 2015. "Holding Local Officials Accountable: School District Performance and School Board Turnover." Paper presented at the 2015 annual meeting of the American Political Science Association.

Ahn, Thomas and Jacob Vigdor. 2013. "Were All Those Standardized Tests for Nothing? The Lessons of No Child Left Behind." A Report Prepared for the National Research Initiative, American Enterprise Institute.

Anzia, Sarah F. 2011. "Election Timing and the Electoral Influence of Interest Groups." Journal of Politics 73:412–427.

Arnold, Douglas R. and Nicholas Carnes. 2012. "Holding Mayors Accountable: New York's Executives from Koch to Bloomberg." American Journal of Political Science 56:949–963.

Atkeson, Lonna Rae and Randall W. Partin. 1995. "Economic and Referendum Voting: A Comparison of Gubernatorial and Senatorial Elections." American Political Science Review 89:99–107.

Berry, Christopher R. and William G. Howell. 2007. "Accountability and Local Elections: Rethinking Retrospective Voting." Journal of Politics 69:844–858.

Brown, Adam R. 2010. "Are Governors Responsible for the State Economy? Partisanship, Blame, and Divided Federalism." Journal of Politics 72:605–615.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." Econometrica 82:2295–2326.

Chakrabarti, Rajashri. 2014. "Incentives and Responses Under No Child Left Behind: Credible Threats and the Role of Competition." Journal of Public Economics 110:124–146.

Charbonneau, Étienne and Greg G. Van Ryzin. 2012. "Performance Measures and Parental Satisfaction with New York City Schools." American Review of Public Administration 42:54–65.

Chingos, Matthew M., Michael Henderson and Martin R. West. 2012. "Citizen Perceptions of Government Service Quality: Evidence from Public Schools." Quarterly Journal of Political Science 7:411–445.

Dee, Thomas S. and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." Journal of Policy Analysis and Management 30:418–446.

Figlio, David N. and Lawrence W. Kenny. 2009. "Public Sector Performance Measurement and Stakeholder Support." Journal of Public Economics 93:1069–1077.

Gelman, Andrew and Guido Imbens. 2014. "Why High-order Polynomials Should not be Used in Regressional Discontinuity Desgns." NBER Working Paper No. 20405.

Hanushek, Eric A. and Margaret E. Raymond. 2005. "Does School Accountability Lead to Improved Student Performance?" Journal of Policy Analysis and Management 24:297–327.

Healy, Andrew and Gabriel S. Lenz. 2014. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." American Journal of Political Science 58:31–47.

Healy, Andrew and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." Annual Review of Political Science 16:285–306.

Henig, Jeffrey R. 2013. The End of Exceptionalism in American Education: The Changing Politics of School Reform. Cambridge: Harvard University Press.

Holbein, John B. Forthcoming. "Left Behind? Citizen Responsiveness to Government Performance Information." American Political Science Review .

Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." Journal of Econometrics 142:615–635.

Jacobsen, Rebecca, Andrew Saultz and Jeffrey W. Snyder. 2013. "When Accountability Strategies Collide: Do Policy Changes That Raise Accountability Standards Also Erode Public Satisfaction?" Education Policy 27:360–389.

Jacobsen, Rebecca, Jeffrey W. Snyder and Andrew Saultz. 2014. "Understanding Satisfaction with Schools: The Role of Expectations." Journal of Public Administration Research and Theory .

Jacobson, Gary C. and Samuel Kernell. 1981. Strategy and Choice in Congressional Elections. New Haven: Yale Unviersity Press.

Kogan, Vladimir, Stéphane Lavertu and Zachary Peskowitz. 2015. "Performance Federalism and Local Democracy: Theory and Evidence from School Levy Elections." American Journal of Political Science .

Krebs, Timothy B. 1998. "The Determinants of Candidates' Vote Share and the Advantages of Incumbency in City Council Elections." American Journal of Political Science 42:921–935.

Ladd, Helen F. and Douglas L. Lauren. 2010. "Status versus Growth: The Distributional Effects of School Accountability Policies." Journal of Policy Analysis and Management 29:426–450.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." Journal of Econometrics 142:675–697.

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." Journal of Economic Literature 48:281–355.

Levitt, Steven D. and Chatherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." Legislative Studies Quarterly 22:45–60.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." Journal of Econometrics 142:698 – 714.

Miller, Michael K. 2013. "For the Win! The Effect of Professional Sports Records on Mayoral Elections." Social Science Quarterly 94:59–78.

Oliver, Eric J. and Shang E. Ha. 2007. "Vote Choice in Suburban Elections." American Political Science Review 101:393–408.

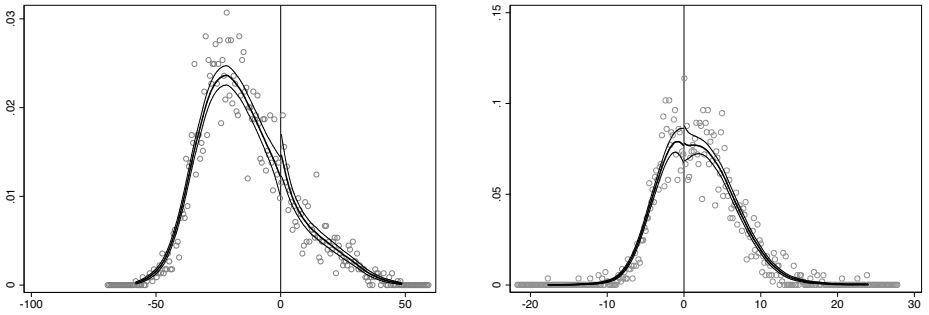
Shaffner, Brian F., Matthew Streb and Gerald Wright. 2001. "Teams Without Uniforms: The Nonpartisan Ballot in State and Local Elections." Political Research Quarterly 54:7–30.

Stone, Walter J., Sarah A. Fulton, Cherie D. Maestas and L. Sandy Maisel. 2010. "Incumbency Reconsidered: Prospects, Strategic Retirement, and Incumbent Quality in U.S. House Elections." Journal of Politics 72:178–190.

APPENDIX

McCrary Test

Figure A.1 presents the McCrary tests for our adjusted running variables. In the absence of manipulation, the distribution of district running variables should be smooth across the threshold that determines which of two designations each district receives. If manipulation exists, on the other hand, we would see significantly more districts just to the right of the threshold compared to immediately to the left of it. Using the adjusted running variables, the figures reveal no significant discontinuities at the cutoffs.



(a) Density of district AYP running variable (adjusted)

(b) Density of running variable for upward adjustment in state designation (using value-added)

Figure A.1: McCrary Tests Using Adjusted Running Variables

Balance Test Plots

As an additional test, we checked covariate balance on each side of the thresholds for a large set of district-level variables, including the measures we excluded from our AYP calculations. Overall, we found no evidence of manipulation, with the number of significant differences roughly the same as one would expect due simply to chance alone.

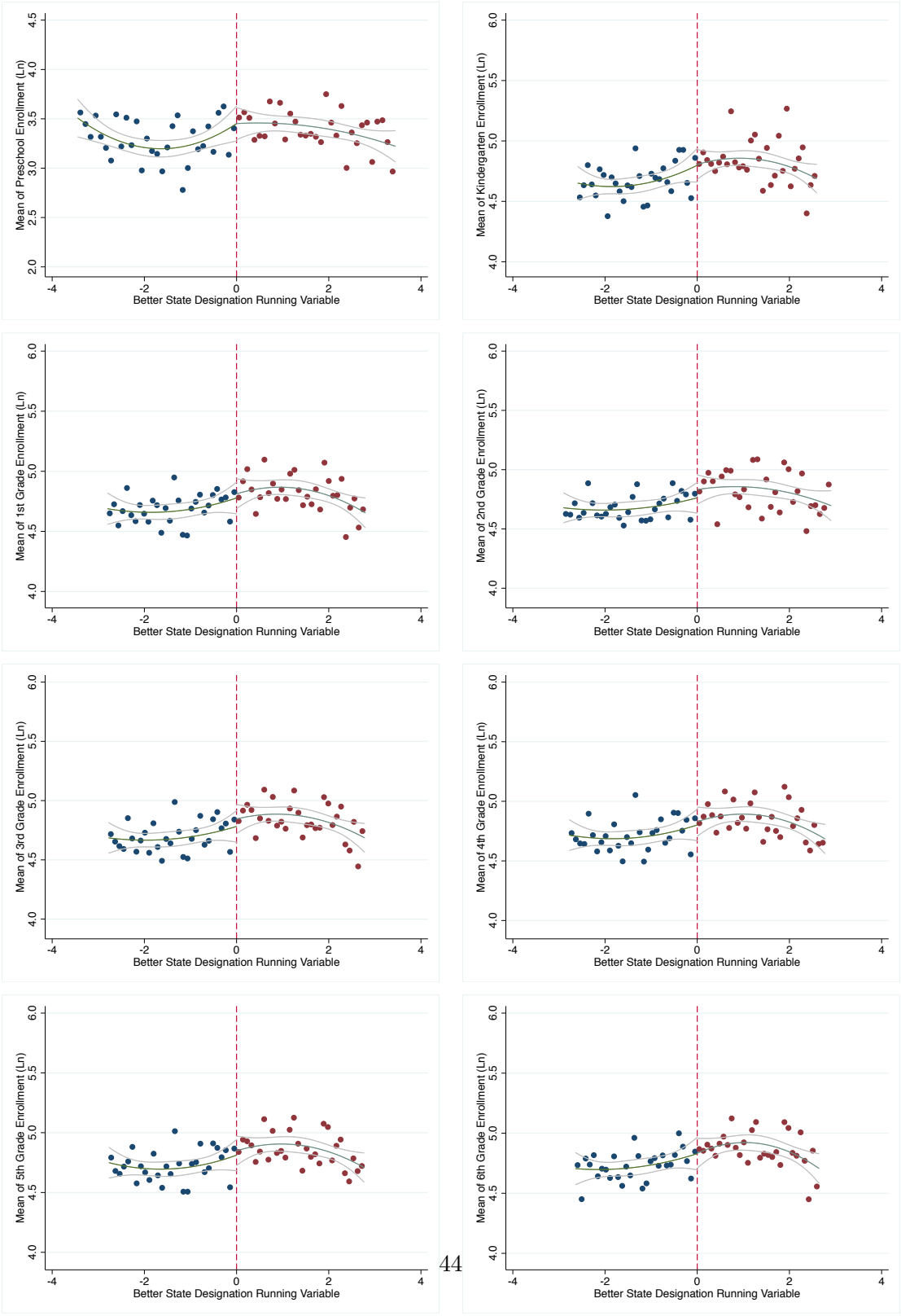

Mean of Preschool Enrollment (Ln)
Mean of Kindergarten Enrollment (Ln)
Mean of 1st Grade Enrollment (Ln)
Mean of 2nd Grade Enrollment (Ln)
Mean of 3rd Grade Enrollment (Ln)
Mean of 4th Grade Enrollment (Ln)
Mean of 5th Grade Enrollment (Ln)
Mean of 6th Grade Enrollment (Ln)
Better State Designation Running Variable

Figure A.2

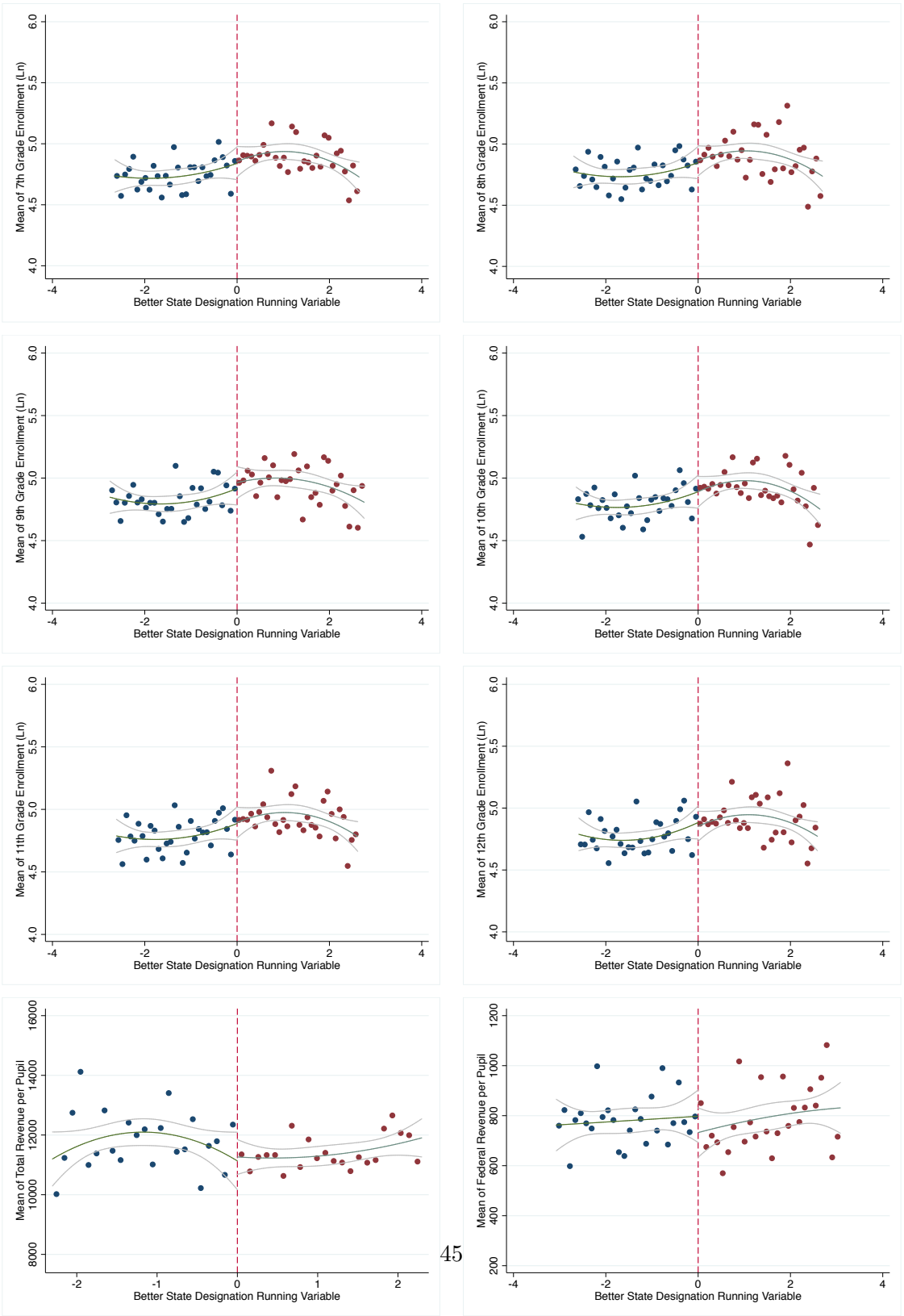

Mean of 7th Grade Enrollment (Ln)
Mean of 8th Grade Enrollment (Ln)
Mean of 9th Grade Enrollment (Ln)
Mean of 10th Grade Enrollment (Ln)
Mean of 11th Grade Enrollment (Ln)
Mean of 12th Grade Enrollment (Ln)
Mean of Total Revenue per Pupil
Mean of Federal Revenue per Pupil
Better State Designation Running Variable

Figure A.3

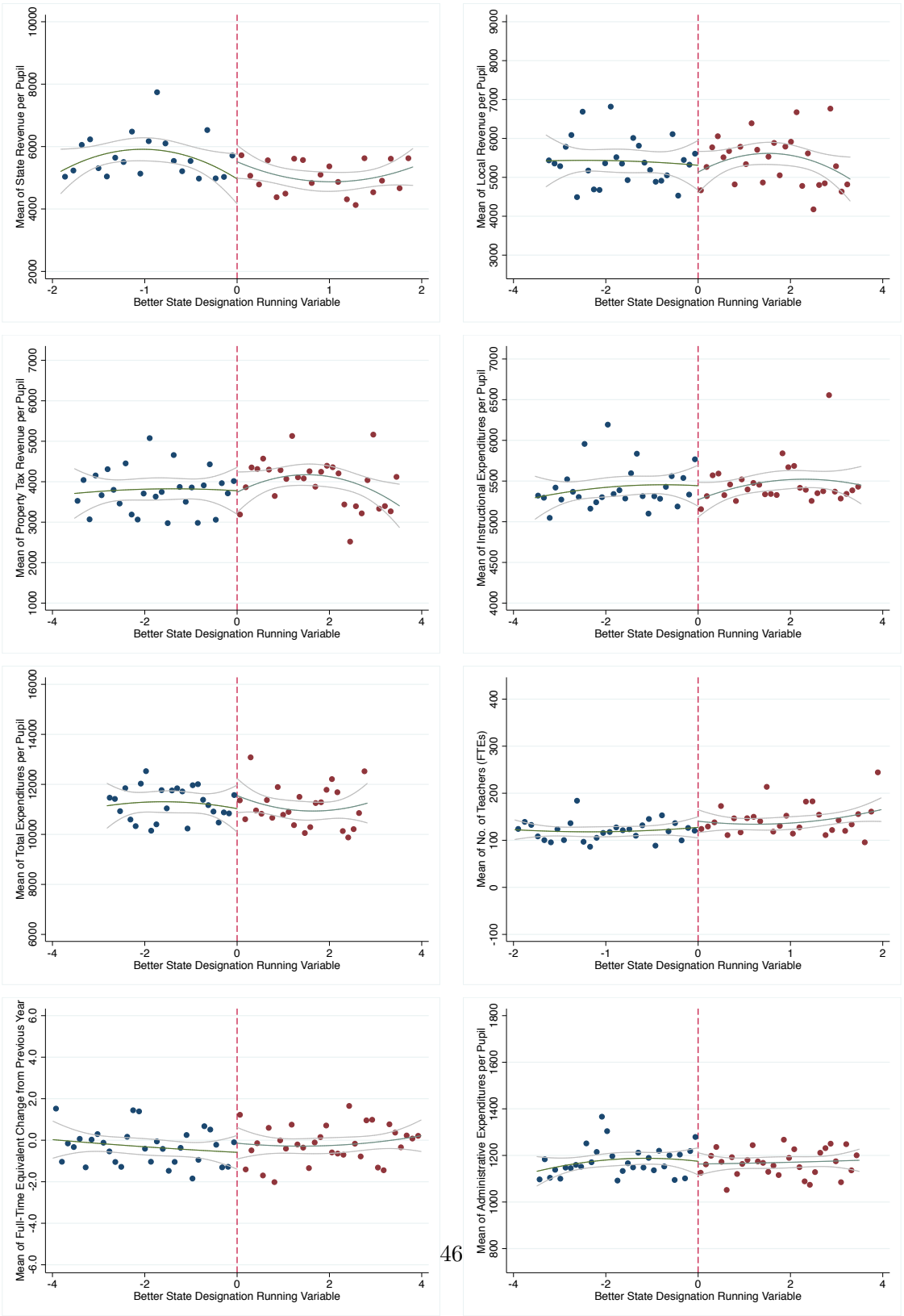

Mean of State Revenue per Pupil
Mean of Local Revenue per Pupil
Mean of Property Tax Revenue per Pupil
Mean of Instructional Expenditures per Pupil
Mean of Total Expenditures per Pupil
Mean of No. of Teachers (FTEs)
Mean of Full-Time Equivalent Change from Previous Year
Mean of Administrative Expenditures per Pupil
Better State Designation Running Variable

Figure A.4

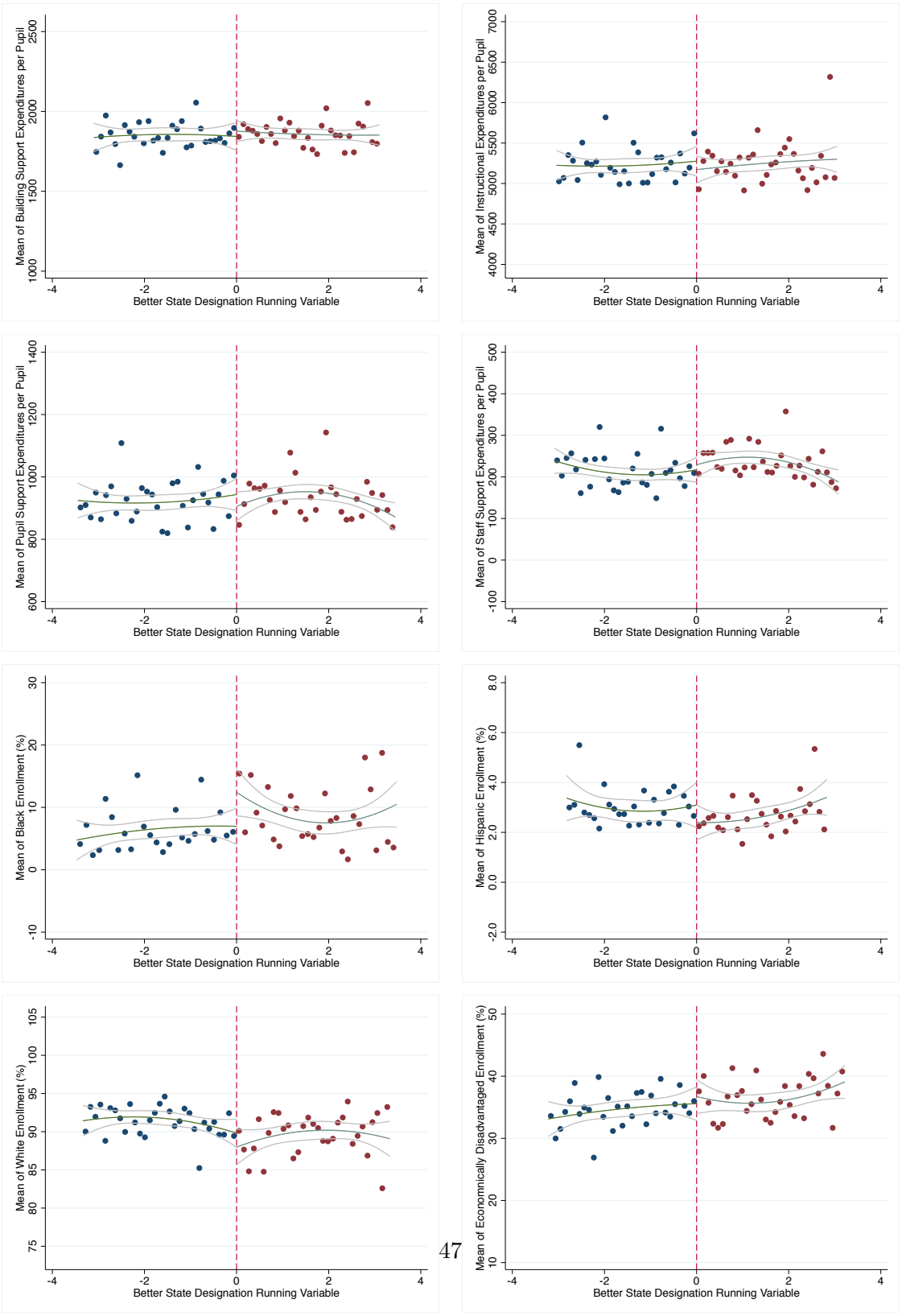

Mean of Building Support Expenditures per Pupil
1000 1500 2000 2500
Mean of Instructional Expenditures per Pupil
4000 4500 5000 5500 6000 6500 7000
Mean of Pupil Support Expenditures per Pupil
600 800 1000 1200 1400
Mean of Staff Support Expenditures per Pupil
-100 0 100 200 300 400 500
Mean of Black Enrollment (%)
-10 0 10 20 30
Mean of Hispanic Enrollment (%)
-2.0 0.0 2.0 4.0 6.0 8.0
Mean of White Enrollment (%)
75 80 85 90 95 100 105
Mean of Economically Disadvantaged Enrollment (%)
10 20 30 40 50
-4 -2 0 2 4
Better State Designation Running Variable

Figure A.5

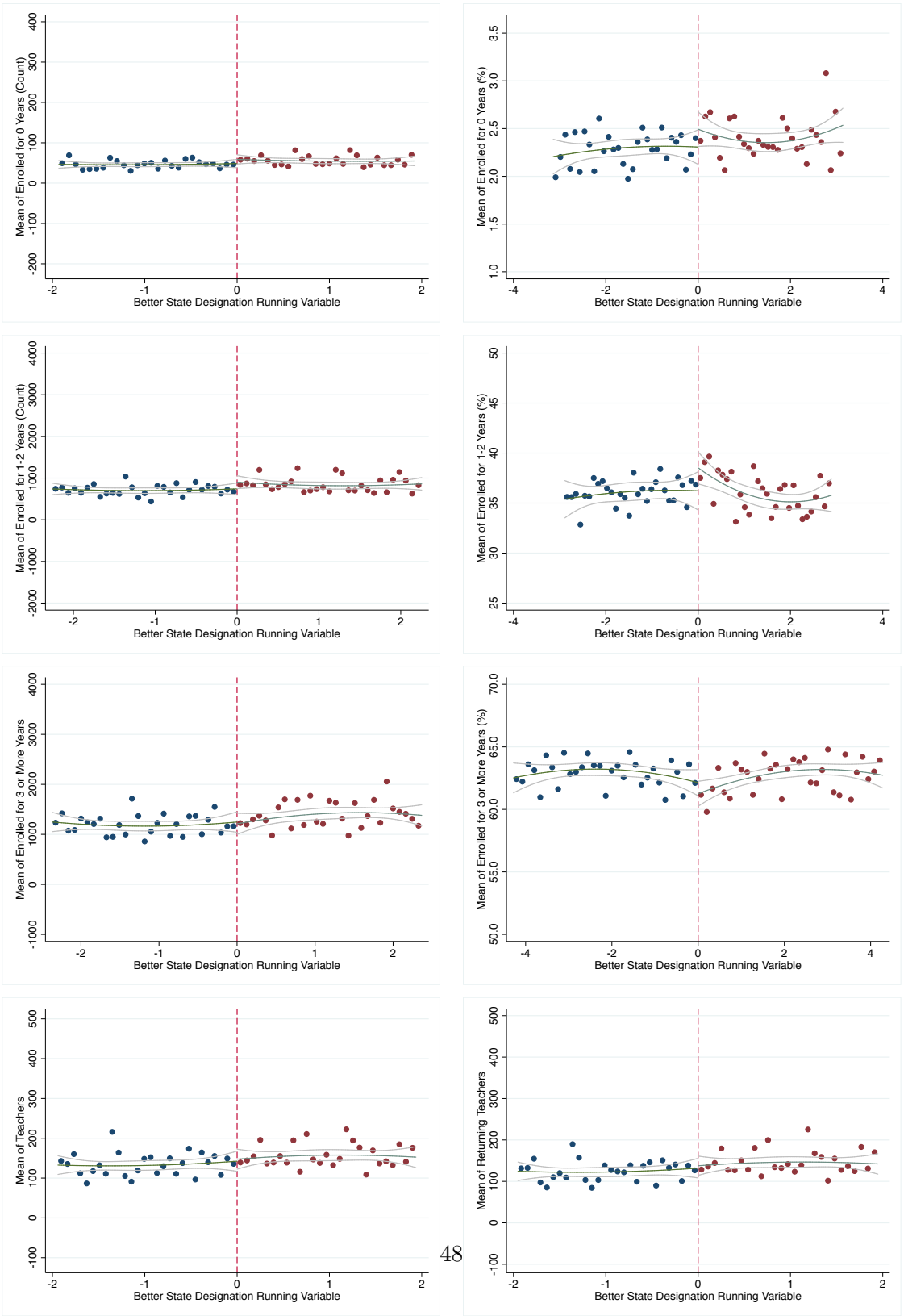

Mean of Enrolled for 0 Years (Count)
Mean of Enrolled for 0 Years (%)
Mean of Enrolled for 1-2 Years (Count)
Mean of Enrolled for 1-2 Years (%)
Mean of Enrolled for 3 or More Years
Mean of Enrolled for 3 or More Years (%)
Mean of Teachers
Mean of Returning Teachers
Better State Designation Running Variable

Figure A.6

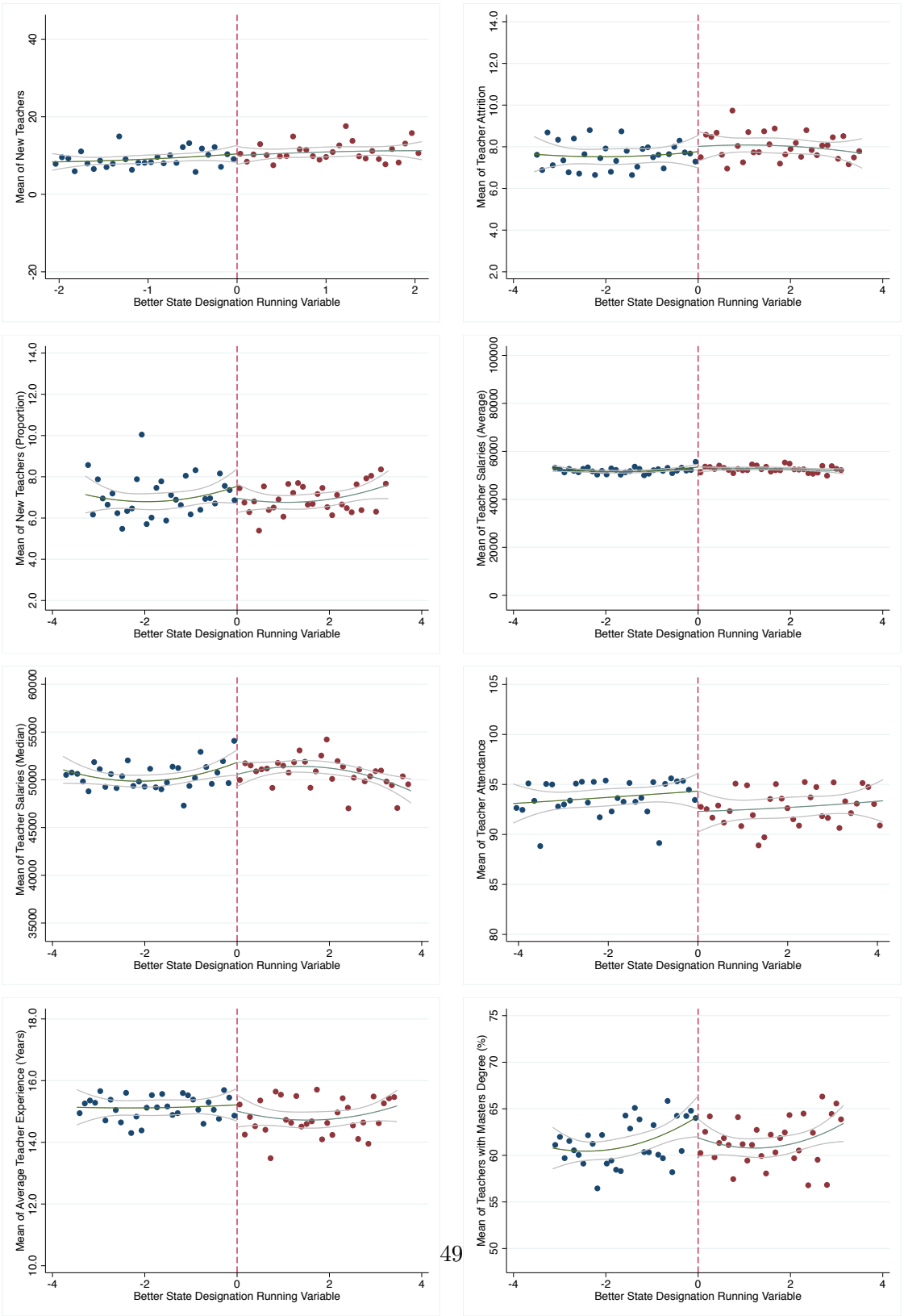

Mean of New Teachers
Better State Designation Running Variable
Mean of Teacher Attrition
Better State Designation Running Variable
Mean of New Teachers (Proportion)
Better State Designation Running Variable
Mean of Teacher Salaries (Average)
Better State Designation Running Variable
Mean of Teacher Salaries (Median)
Better State Designation Running Variable
Mean of Teacher Attendance
Better State Designation Running Variable
Mean of Average Teacher Experience (Years)
Better State Designation Running Variable
Mean of Teachers with Masters Degree (%)
Better State Designation Running Variable

Figure A.7

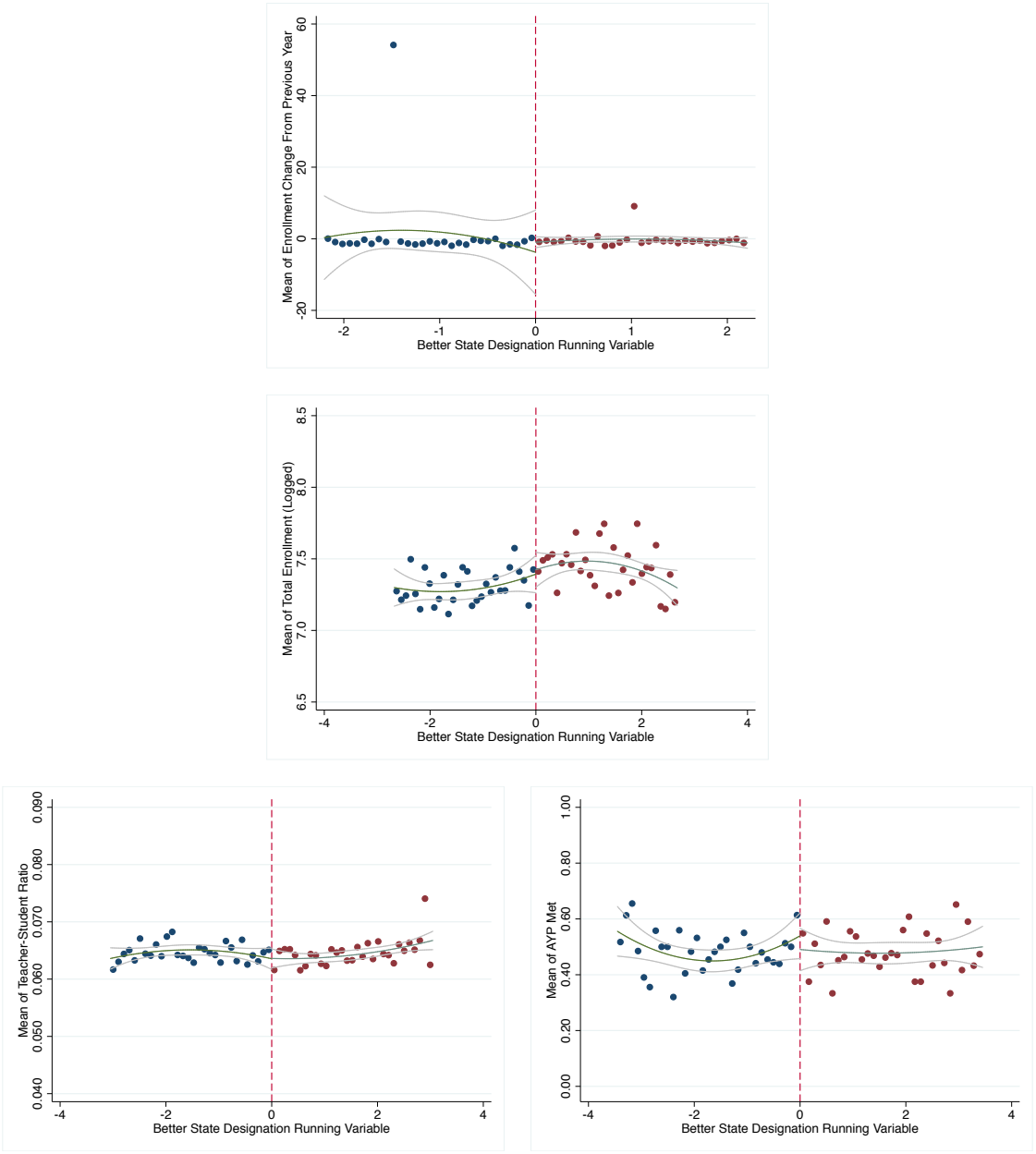

Mean of Enrollment Change From Previous Year
-20
0
20
40
60
-2
-1
0
1
2
Better State Designation Running Variable
Mean of Total Enrollment (Logged)
6.5
7.0
7.5
8.0
8.5
-4
-2
0
2
4
Better State Designation Running Variable
Mean of Teacher-Student Ratio
0.040
0.050
0.060
0.070
0.080
0.090
-4
-2
0
2
4
Better State Designation Running Variable
Mean of AYP Met
0.00
0.20
0.40
0.60
0.80
1.00
-4
-2
0
2
4
Better State Designation Running Variable

Figure A.8

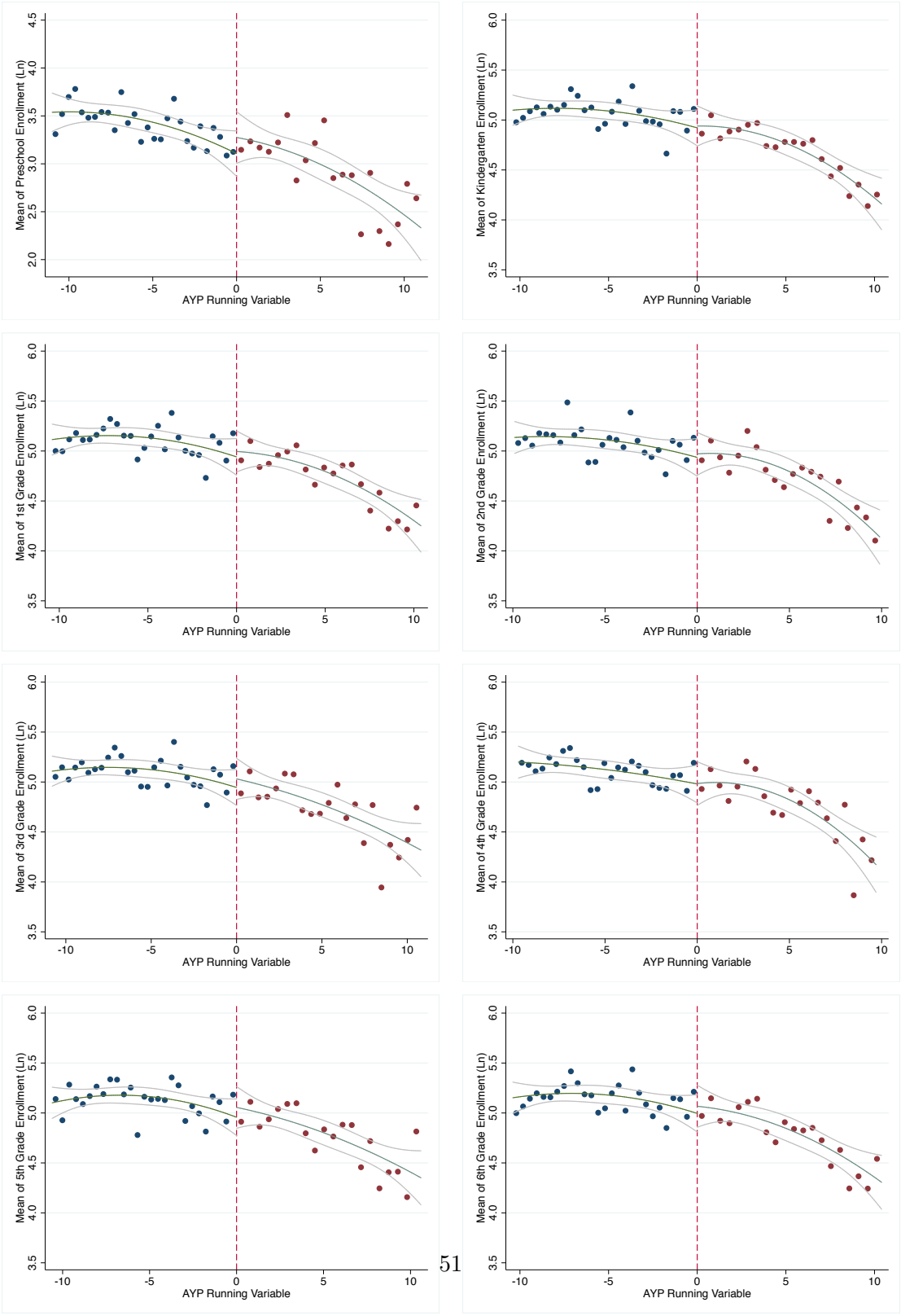

Mean of Preschool Enrollment (Ln)
Mean of Kindergarten Enrollment (Ln)
Mean of 1st Grade Enrollment (Ln)
Mean of 2nd Grade Enrollment (Ln)
Mean of 3rd Grade Enrollment (Ln)
Mean of 4th Grade Enrollment (Ln)
Mean of 5th Grade Enrollment (Ln)
Mean of 6th Grade Enrollment (Ln)
AYP Running Variable

Figure A.9

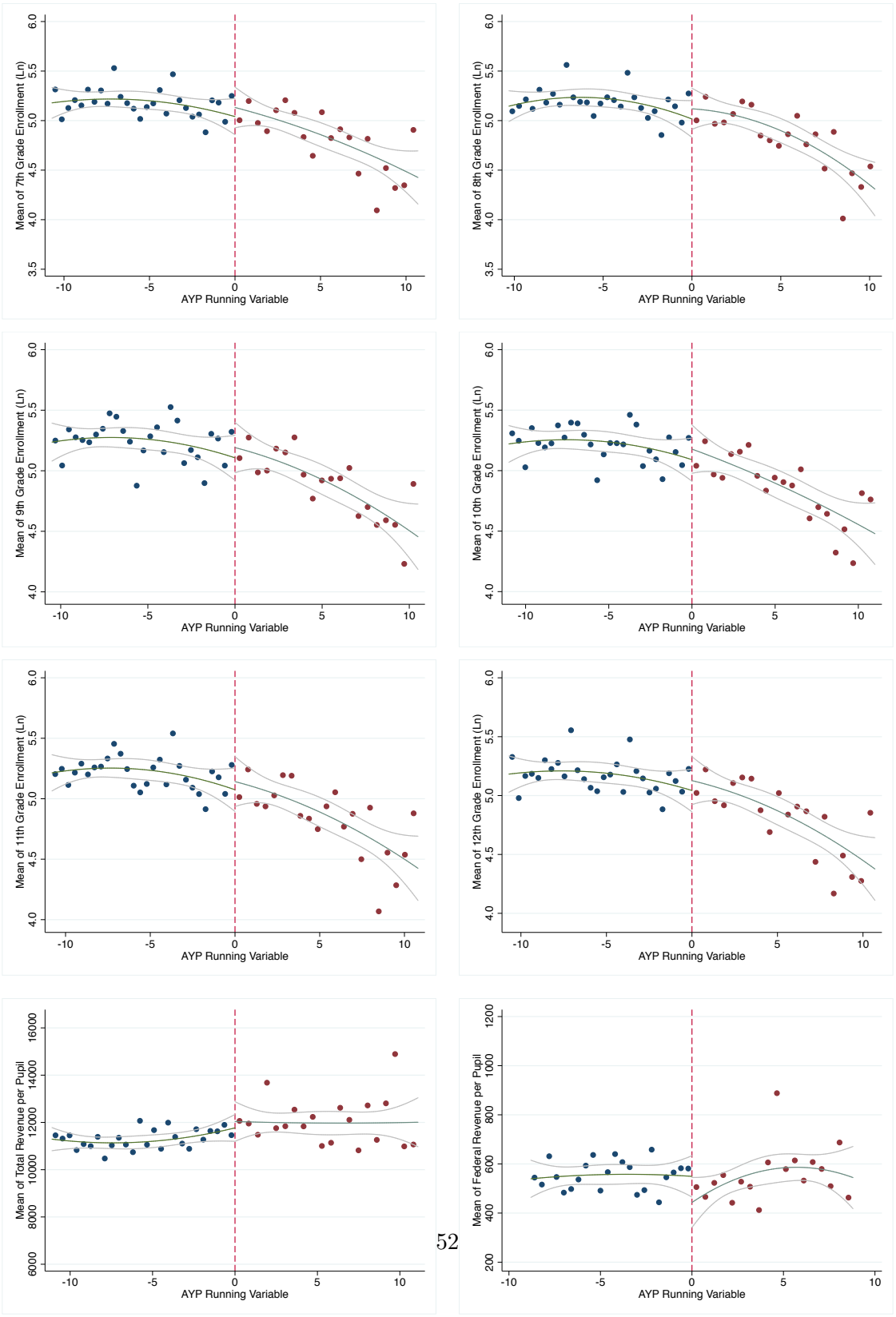

Mean of 7th Grade Enrollment (Ln)
Mean of 8th Grade Enrollment (Ln)
Mean of 9th Grade Enrollment (Ln)
Mean of 10th Grade Enrollment (Ln)
Mean of 11th Grade Enrollment (Ln)
Mean of 12th Grade Enrollment (Ln)
Mean of Total Revenue per Pupil
Mean of Federal Revenue per Pupil
AYP Running Variable

Figure A.10

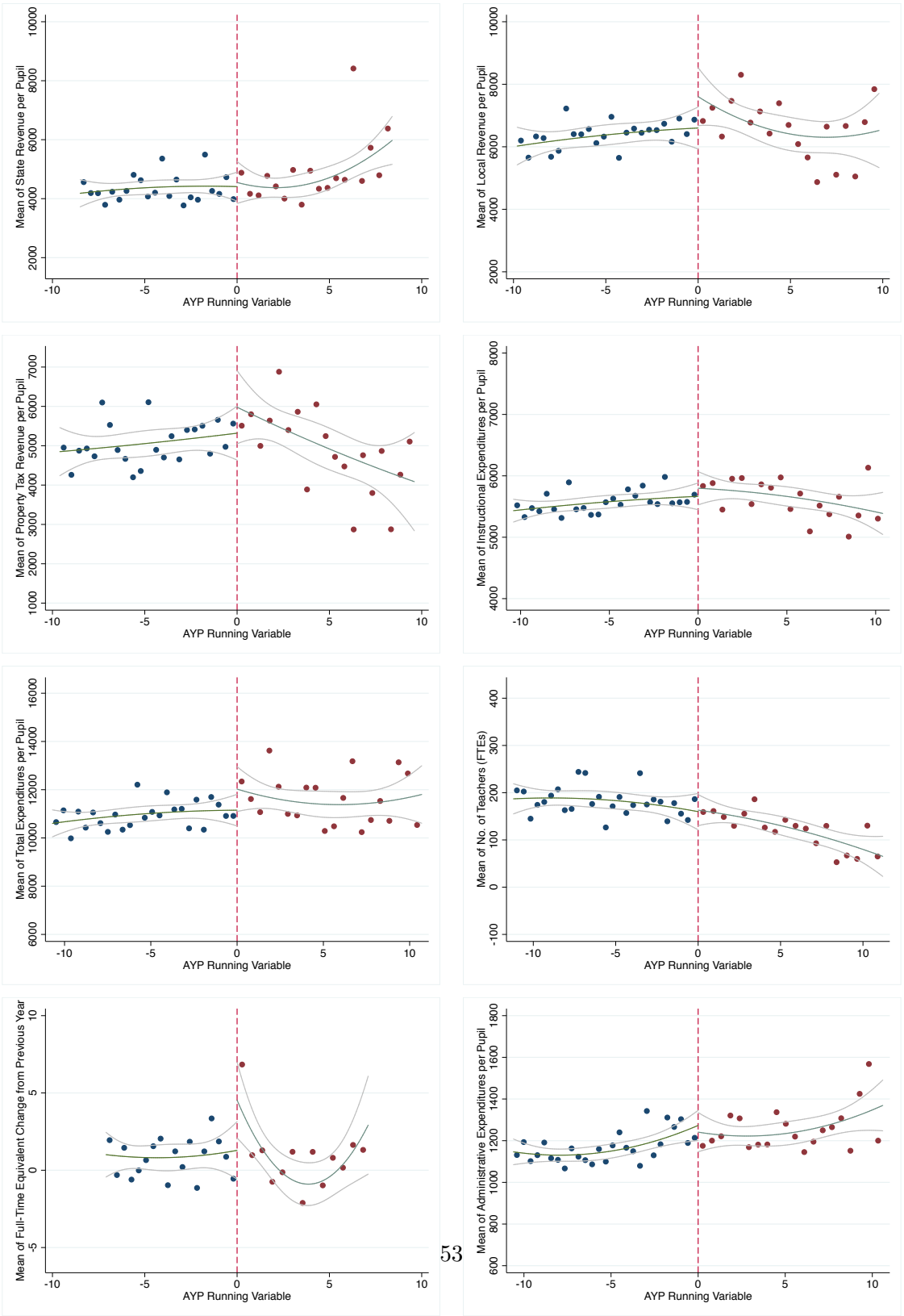

Mean of State Revenue per Pupil
Mean of Local Revenue per Pupil
Mean of Property Tax Revenue per Pupil
Mean of Instructional Expenditures per Pupil
Mean of Total Expenditures per Pupil
Mean of No. of Teachers (FTEs)
Mean of Full-Time Equivalent Change from Previous Year
Mean of Administrative Expenditures per Pupil
AYP Running Variable

Figure A.11

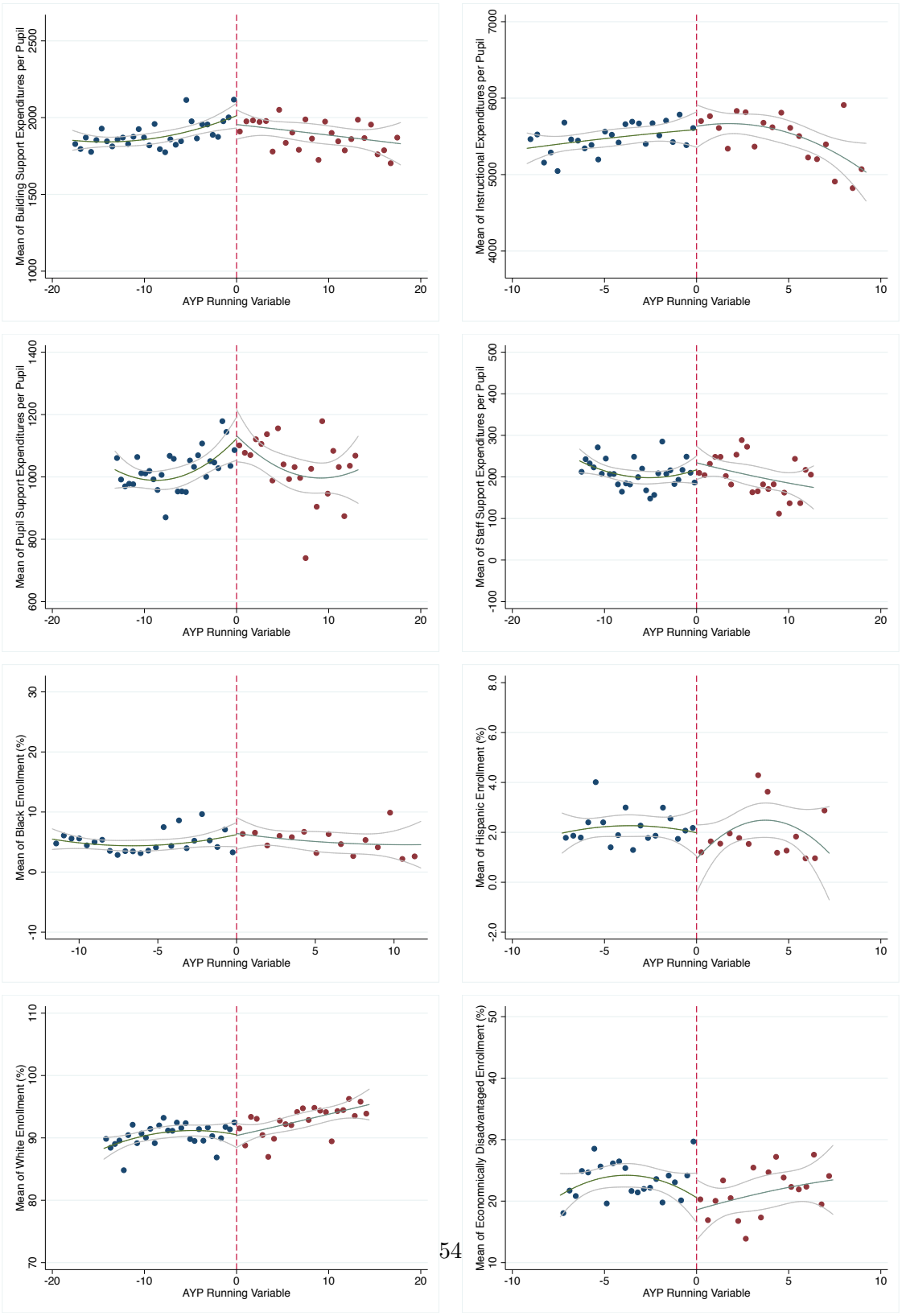

Mean of Building Support Expenditures per Pupil
1000
1500
2000
2500
-20
-10
0
10
20
AYP Running Variable
Mean of Instructional Expenditures per Pupil
4000
5000
6000
7000
-10
-5
0
5
10
AYP Running Variable
Mean of Pupil Support Expenditures per Pupil
600
800
1000
1200
1400
-20
-10
0
10
20
AYP Running Variable
Mean of Staff Support Expenditures per Pupil
-100
0
100
200
300
400
500
-20
-10
0
10
20
AYP Running Variable
Mean of Black Enrollment (%)
-10
0
10
20
30
-10
-5
0
5
10
AYP Running Variable
Mean of Hispanic Enrollment (%)
-2.0
0.0
2.0
4.0
6.0
8.0
-10
-5
0
5
10
AYP Running Variable
Mean of White Enrollment (%)
70
80
90
100
110
-20
-10
0
10
20
AYP Running Variable
Mean of Economically Disadvantaged Enrollment (%)
10
20
30
40
50
-10
-5
0
5
10
AYP Running Variable

Figure A.12

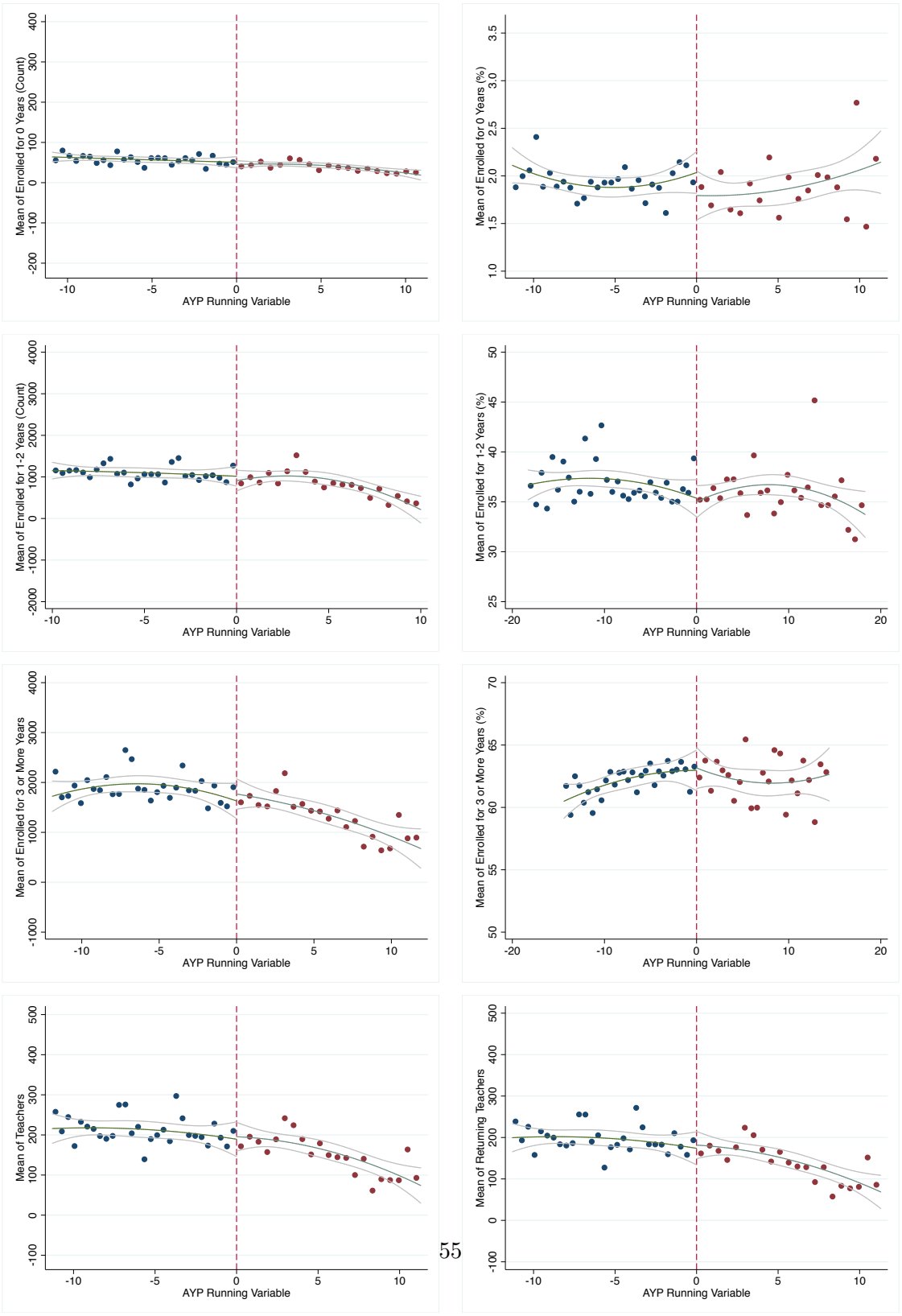

Mean of Enrolled for 0 Years (Count)
AYP Running Variable
Mean of Enrolled for 0 Years (%)
AYP Running Variable
Mean of Enrolled for 1-2 Years (Count)
AYP Running Variable
Mean of Enrolled for 1-2 Years (%)
AYP Running Variable
Mean of Enrolled for 3 or More Years
AYP Running Variable
Mean of Enrolled for 3 or More Years (%)
AYP Running Variable
Mean of Teachers
AYP Running Variable
Mean of Returning Teachers
AYP Running Variable

Figure A.13

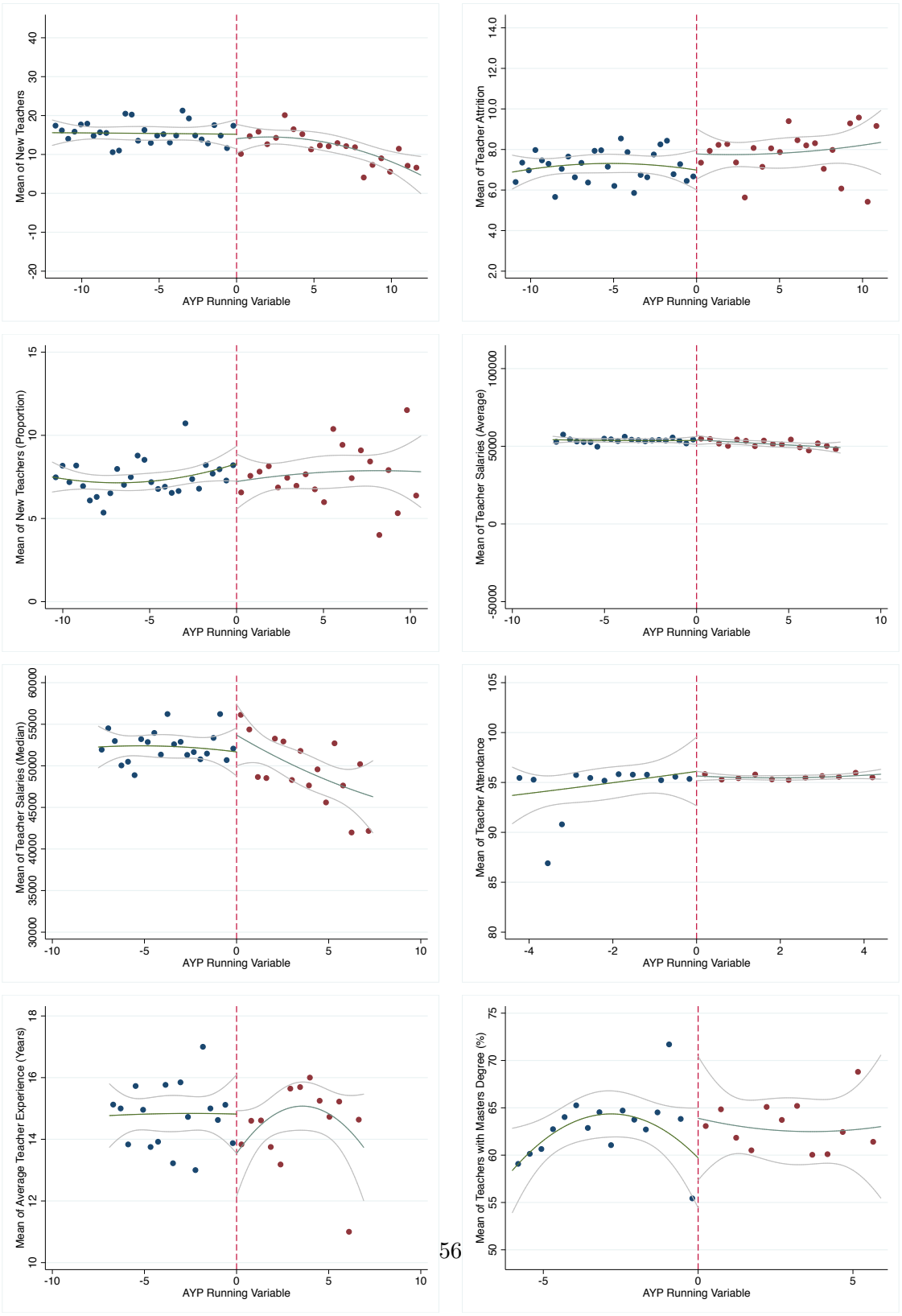

Mean of New Teachers
AYP Running Variable
Mean of Teacher Attrition
AYP Running Variable
Mean of New Teachers (Proportion)
AYP Running Variable
Mean of Teacher Salaries (Average)
AYP Running Variable
Mean of Teacher Salaries (Median)
AYP Running Variable
Mean of Teacher Attendance
AYP Running Variable
Mean of Average Teacher Experience (Years)
AYP Running Variable
Mean of Teachers with Masters Degree (%)
AYP Running Variable

Figure A.14

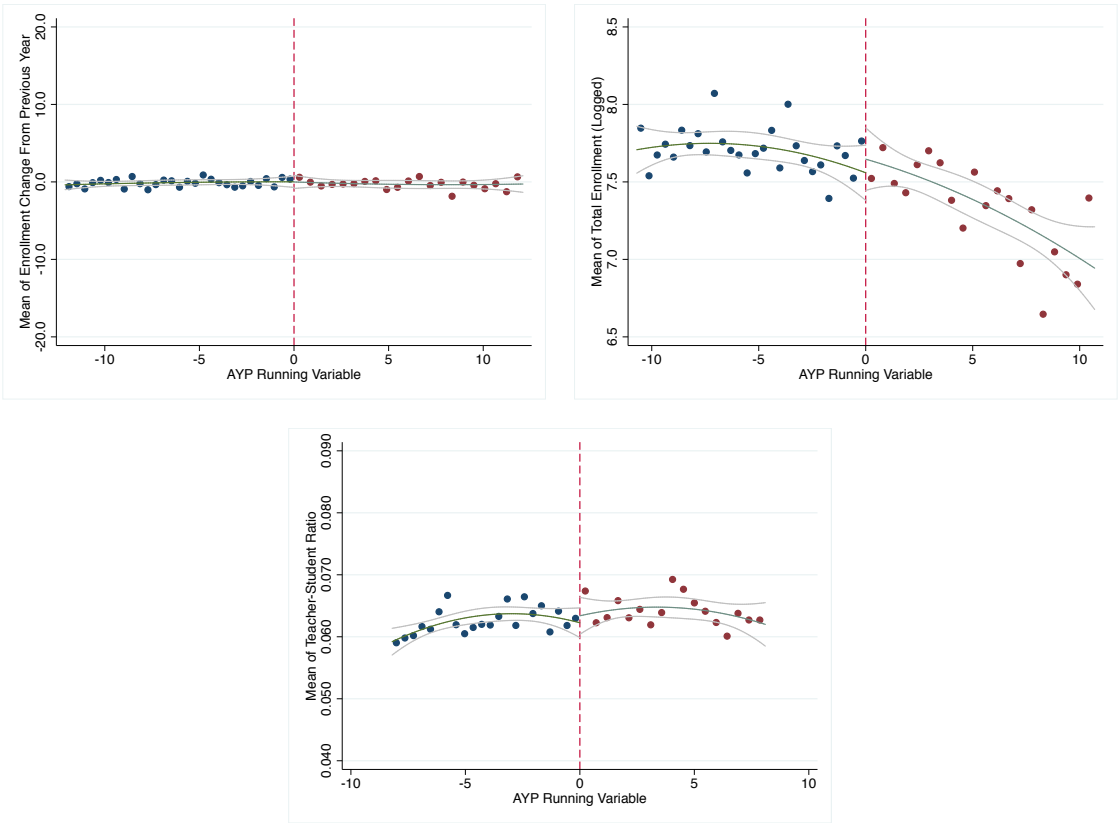
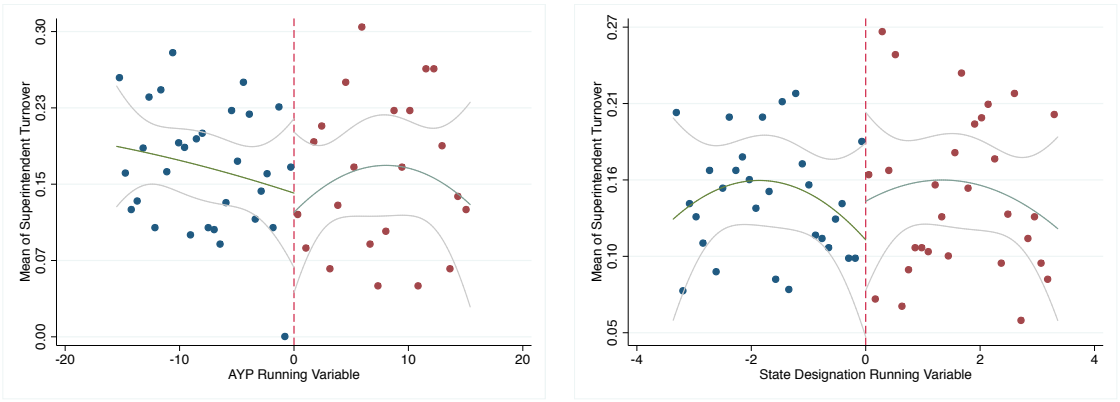

Mean of Enrollment Change From Previous Year
20.0
10.0
0.0
-10.0
-20.0
-10
-5
0
5
10
AYP Running Variable
Mean of Total Enrollment (Logged)
8.5
8.0
7.5
7.0
6.5
-10
-5
0
5
10
AYP Running Variable
Mean of Teacher-Student Ratio
0.090
0.080
0.070
0.060
0.050
0.040
-10
-5
0
5
10
AYP Running Variable

Figure A.15

Graphical RDD Plots for Superintendent Turnover



(a) Federal Adequate Yearly Progress running variable and superintendent turnover

(b) "Better state designation" running variable and superintendent turnover

Figure A.16: The figure displays the relationship between the performance information running variables and superintendent turnover. The circles indicate the local average school board turnover; the dark lines represent the quadratic best fit estimated separately to the left and the right of the threshold; and the 95 percent confidence interval is demarcated using light lines.

School-Level Designations and Board Turnover

Table A.1: District Performance Information on School Board Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.00490	0.00884	-0.0135	0.0192
	(0.0162)	(0.0204)	(0.0192)	(0.0246)
Academic Emergency	-0.115	-0.166	-0.0430	-0.0872
	(0.0731)	(0.122)	(0.189)	(0.344)
Continuous Improvement	-0.0254	0.0505	0.0624	0.157^{+}
	(0.0478)	(0.0570)	(0.0670)	(0.0856)
Effective	-0.0723	0.0277	0.0355	0.128
	(0.0560)	(0.0675)	(0.0709)	(0.0925)
Excellent	-0.0549	0.0408	0.0696	0.169^{+}
	(0.0604)	(0.0726)	(0.0754)	(0.0987)
Excellent w/ Distinction	-0.0865	-0.0398	0.0596	0.154
	(0.0687)	(0.0817)	(0.0837)	(0.109)
Effective and Above (% of Schools)	-0.000403	0.0000200	-0.000926^{+}	0.0000447
	(0.000471)	(0.000578)	(0.000526)	(0.000747)
AYP Met (% of Schools)	0.0000719	0.0000849	0.0000966	0.000170
	(0.000385)	(0.000450)	(0.000389)	(0.000516)
Observations	7313	7313	5910	5910
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: District Performance Information on School Board Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.00474	0.00962	-0.0145	0.0210
	(0.0163)	(0.0207)	(0.0193)	(0.0248)
Academic Emergency	-0.115	-0.155	-0.0795	-0.309
	(0.0713)	(0.132)	(0.171)	(0.244)
Continuous Improvement	-0.0308	0.0291	0.0653	0.174^{+}
	(0.0489)	(0.0596)	(0.0698)	(0.0910)
Effective	-0.0859	-0.00403	0.0344	0.145
	(0.0573)	(0.0699)	(0.0733)	(0.0964)
Excellent	-0.0638	0.0170	0.0657	0.183^{+}
	(0.0611)	(0.0745)	(0.0776)	(0.102)
Excellent w/ Distinction	-0.0973	-0.0579	0.0522	0.167
	(0.0697)	(0.0832)	(0.0857)	(0.112)
Effective and Above (% of Enrollment)	-0.000207	0.00000696	-0.000855	-0.000222
	(0.000473)	(0.000572)	(0.000528)	(0.000729)
AYP Met (% of Enrollment)	0.0000228	0.000268	0.0000166	-0.0000796
	(0.000324)	(0.000376)	(0.000343)	(0.000461)
Observations	7190	7190	5861	5861
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

State RD Interactions

Table A.3: One-Unit Increase in State Designation on School Board Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
Increase to Academic Watch	-0.152	0.194	-0.129	0.0556
	(0.0988)	(0.141)	(0.14)	(0.172)
Increase to Continuous Improvement	-0.068	0.056	-0.065	0.1
	(0.107)	(0.16)	(0.114)	(0.155)
Increase to Effective	0.1	0.085	-0.048	-0.056
	(0.066)	(0.083)	(0.089)	(0.131)
Increase to Excellent	0.067^{+}	0.06	0.027	0.036
	(0.037)	(0.06)	(0.039)	(0.063)
Increase to Excellent w/ Distinct.	0.007	-0.029	-0.046	-0.051
	(0.043)	(0.065)	(0.045)	(0.065)
Observations	5524	3140	3823	2531
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The table reports linear combinations of the RD coefficients and the interaction effects for each state designation.
The unit of analysis in the models is individual incumbent school board members with
performance information measured at the district level. The analysis includes years 2008-2012.
We report heteroskedasticity robust standard errors clustered at district-year in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: One-Unit Increase in State Designation on School Board Vote Share (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
Increase to Academic Watch	-0.00831	-0.0812	-0.0865	-0.0878
	(0.0643)	(0.122)	(0.0726)	(0.0895)
Increase to Continuous Improvement	0.037	0.049	0.086	0.234*
	(0.061)	(0.101)	(0.059)	(0.11)
Increase to Effective	0.042	0.082	0.014	0.018
	(0.072)	(0.1)	(0.094)	(0.117)
Increase to Excellent	-0.046+	-0.012	-0.01	-0.015
	(0.028)	(0.058)	(0.034)	(0.051)
Increase to Excellent w/ Distinct.	0.038	-0.01	0.069*	0.086+
	(0.03)	(0.061)	(0.031)	(0.05)
Observations	1195	355	833	431
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The table reports linear combinations of the RD coefficients and the interaction effects for each state designation. The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: One-Unit Increase in State Designation on Superintendent Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
Increase to Academic Watch	0.286*	0.177	0.362**	0.468	0.645***	0.49*
	(0.11)	(0.144)	(0.136)	(0.286)	(0.184)	(0.232)
Increase to Continuous Improvement	0.088	0.189	-0.065	0.079	0.006***	-0.061*
	(0.086)	(0.124)	(0.116)	(0.167)	(0.16)	(0.29)
Increase to Effective	-0.086	-0.074	0.045	0.1	0.107^{+}	0.104
	(0.082)	(0.103)	(0.071)	(0.097)	(0.062)	(0.081)
Increase to Excellent	0.015	0.03	0.021	0.039	0.005	0.02
	(0.03)	(0.051)	(0.034)	(0.066)	(0.04)	(0.061)
Increase to Excellent w/ Distinct.	0.011	0.014	0.004	0.023	0.034	0.055
	(0.033)	(0.054)	(0.039)	(0.067)	(0.049)	(0.071)
Observations	3727	1947	2618	1132	1539	1000
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes	No	Yes

The table reports linear combinations of the RD coefficients and the interaction effects for each state designation.
The unit of analysis in the models is individual superintendents with
performance information measured at the district level. The analysis includes years 2008-2012.
We report heteroskedasticity robust standard errors clustered at district-year in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Incumbent Vote Share Results

Table A.6: District Performance Information on School Board Incumbent Vote Share (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	-0.00795	-0.00296	0.00592	0.000626
	(0.0131)	(0.0192)	(0.0147)	(0.0279)
Academic Emergency	-0.0695^{+}	-0.0667	-0.0375	0.00850
	(0.0390)	(0.0743)	(0.0664)	(0.191)
Continuous Improvement	0.0428	0.0604	0.0296	0.0724
	(0.0392)	(0.0748)	(0.0381)	(0.0832)
Effective	0.0503	0.0864	0.0445	0.0606
	(0.0403)	(0.0776)	(0.0429)	(0.0972)
Excellent	0.0269	0.0721	0.0347	0.0740
	(0.0428)	(0.0804)	(0.0449)	(0.104)
Excellent w/ Distinction	0.0411	0.0780	0.0543	0.0992
	(0.0488)	(0.0803)	(0.0513)	(0.107)
Observations	1475	1475	1272	1272
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: District AYP Passage on School Board Vote Share (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	-0.0114	-0.0340	0.0415	0.0773
	(0.0876)	(0.0879)	(0.0578)	(0.0962)
Observations	1272	236	1272	402
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

Heteroskedasticity robust standard errors clustered at district level

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.8: Better State Designation on School Board Vote Share (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
Better State Desig.	0.00118	-0.00501	0.0320	0.0266
	(0.0230)	(0.0571)	(0.0264)	(0.0516)
Observations	1195	355	833	431
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year level in parentheses.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Controlling for Proficiency

Table A.9: District Performance Information on School Board Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.00644	-0.0133	-0.0112	0.0232
	(0.0186)	(0.0242)	(0.0192)	(0.0246)
Academic Emergency	-0.0864	-0.286	-0.0398	-0.0867
	(0.0680)	(0.204)	(0.189)	(0.353)
Continuous Improvement	0.0150	-0.0363	0.0707	0.136
	(0.0776)	(0.101)	(0.0671)	(0.0861)
Effective	-0.0579	-0.119	0.0386	0.106
	(0.0841)	(0.112)	(0.0708)	(0.0940)
Excellent	-0.0432	-0.103	0.0770	0.149
	(0.0880)	(0.114)	(0.0754)	(0.100)
Excellent w/ Distinction	-0.0725	-0.174	0.0741	0.133
	(0.0957)	(0.121)	(0.0834)	(0.111)
% Prof. in Math	-0.000890	0.00618^{+}	-0.00182	0.00642
	(0.00279)	(0.00354)	(0.00276)	(0.00417)
% Prof. in Reading	-0.00205	0.00506	-0.00344	-0.00335
	(0.00389)	(0.00415)	(0.00376)	(0.00543)
Observations	5910	5910	5910	5910
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.10: District Performance Information on School Board Incumbent Vote Share (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	-0.0150	-0.00649	0.00649	-0.00113
	(0.0153)	(0.0255)	(0.0147)	(0.0273)
Academic Emergency	-0.0251	-0.234*	-0.0456	-0.0180
	(0.0636)	(0.108)	(0.0627)	(0.189)
Continuous Improvement	0.0523	0.188	0.0313	0.101
	(0.0607)	(0.135)	(0.0382)	(0.0862)
Effective	0.0615	0.201	0.0427	0.0845
	(0.0615)	(0.140)	(0.0439)	(0.100)
Excellent	0.0317	0.185	0.0304	0.100
	(0.0649)	(0.146)	(0.0456)	(0.107)
Excellent w/ Distinction	0.0426	0.173	0.0483	0.120
	(0.0692)	(0.147)	(0.0521)	(0.109)
% Prof. in Math	-0.0000200	-0.00258	-0.000592	-0.00273
	(0.00235)	(0.00454)	(0.00217)	(0.00453)
% Prof. in Reading	0.00314	0.00882	0.00280	0.00931
	(0.00294)	(0.00570)	(0.00298)	(0.00599)
Observations	1272	1272	1272	1272
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.11: Performance Information on Superintendent Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	0.0163	0.0136	0.0140	0.0113	0.00222	-0.0175
	(0.0141)	(0.0176)	(0.0145)	(0.0178)	(0.0154)	(0.0206)
Academic Emergency	0.313^{+}	0.496***	-0.0206	0.0439	0.0384	0.0761
	(0.185)	(0.134)	(0.121)	(0.154)	(0.109)	(0.123)
Continuous Improvement	0.0934	0.0531	0.0527	0.000812	0.0930*	0.0454
	(0.0590)	(0.0736)	(0.0539)	(0.0680)	(0.0473)	(0.0724)
Effective	0.0979	0.0569	0.0613	0.0168	0.0753	0.0292
	(0.0667)	(0.0834)	(0.0582)	(0.0735)	(0.0529)	(0.0803)
Excellent	0.116	0.0992	0.0619	0.0353	0.0920	0.0507
	(0.0713)	(0.0884)	(0.0620)	(0.0779)	(0.0580)	(0.0857)
Excellent w/ Distinction	0.115	0.0748	0.0710	0.0281	0.115^{+}	0.0833
	(0.0758)	(0.0931)	(0.0685)	(0.0850)	(0.0689)	(0.101)
% Prof. in Math	-0.00176	-0.00149	-0.00156	-0.000901	-0.00101	-0.000342
	(0.00201)	(0.00262)	(0.00196)	(0.00255)	(0.00218)	(0.00290)
% Prof. in Reading	-0.000923	0.000627	-0.000138	0.00151	0.00151	0.000242
	(0.00275)	(0.00363)	(0.00272)	(0.00359)	(0.00323)	(0.00442)
Observations	4322	4322	4322	4322	3997	3997
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Large vs. Small Districts

Table A.12: District Performance Information on School Board Turnover (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.0202	0.0343	-0.0113	0.0222
	(0.0239)	(0.0294)	(0.0195)	(0.0246)
Academic Emergency	-0.160*	-0.235^{+}	-0.0590	-0.185
	(0.0768)	(0.126)	(0.230)	(0.384)
Continuous Improvement	-0.0386	0.0612	0.0575	0.160^{+}
	(0.0505)	(0.0620)	(0.0634)	(0.0870)
Effective	-0.0826	0.0704	0.0226	0.142
	(0.0598)	(0.0774)	(0.0683)	(0.0939)
Excellent	-0.0887	0.0355	0.0471	0.174^{+}
	(0.0653)	(0.0842)	(0.0736)	(0.100)
Excellent w/ Distinction	-0.110	-0.0388	0.0488	0.168
	(0.0764)	(0.0978)	(0.0821)	(0.111)
Small*AYP Met	-0.0293	-0.0480	-0.00846	-0.00424
	(0.0319)	(0.0387)	(0.0249)	(0.0314)
Small*Academic Emergency	0.259^{+}	0.592**	0.0714	0.652
	(0.144)	(0.201)	(0.253)	(0.451)
Small*Continuous Improvement	0.0373	-0.0541	0.242	0.00352
	(0.114)	(0.131)	(0.212)	(0.388)
Small*Effective	0.0145	-0.104	0.183	-0.0703
	(0.117)	(0.142)	(0.214)	(0.387)
Small*Excellent	0.0586	-0.00748	0.237	0.0214
	(0.121)	(0.152)	(0.215)	(0.387)
Small*Excellent w/ Distinction	0.0330	-0.0278	0.176	-0.0543
	(0.134)	(0.169)	(0.222)	(0.392)
Observations	7311	7311	5903	5903
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.13: District AYP Passage on School Board Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.178	-0.0657	-0.0858	0.193
	(0.125)	(0.140)	(0.131)	(0.175)
Small*AYP Met	-0.0340	0.00855	0.00984	0.0518
	(0.0508)	(0.0616)	(0.0439)	(0.0575)
Observations	5903	1336	5903	1429
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. Heteroskedasticity robust standard errors clustered at district level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.14: Better State Designation on School Board Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
Better State Desig.	0.0445	0.0371	-0.0385	-0.0299
	(0.0350)	(0.0588)	(0.0380)	(0.0597)
Small*BSD	0.0179	-0.00231	0.0492	0.0410
	(0.0226)	(0.0322)	(0.0303)	(0.0376)
Observations	5514	3130	3819	2527
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Superintendent Results For Sample With More than One Year of Tenure

Table A.15: Performance Information on Turnover of Superintendents With More than One Year of Tenure (FE)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	-0.00586	0.000608	-0.00678	-0.000609	-0.00498	-0.00936
	(0.0109)	(0.0144)	(0.0111)	(0.0145)	(0.0122)	(0.0167)
Academic Emergency	0.219	0.239	-0.0778	-0.181^{+}	0.103	0.0624
	(0.154)	(0.223)	(0.0572)	(0.0999)	(0.0997)	(0.118)
Continuous Improvement	-0.00768	-0.0172	0.0302	-0.0159	0.0590^{+}	0.0236
	(0.0339)	(0.0563)	(0.0278)	(0.0389)	(0.0342)	(0.0485)
Effective	-0.000581	0.0149	0.0377	0.0181	0.0305	0.00563
	(0.0389)	(0.0623)	(0.0323)	(0.0442)	(0.0391)	(0.0569)
Excellent	-0.0147	0.0149	0.0225	0.0117	0.0411	0.0304
	(0.0420)	(0.0649)	(0.0356)	(0.0484)	(0.0435)	(0.0625)
Excellent w/ Distinction	-0.0109	0.0119	0.0265	0.00753	0.0445	0.0377
	(0.0464)	(0.0684)	(0.0431)	(0.0544)	(0.0533)	(0.0746)
Observations	4096	4096	4096	4096	3578	3578
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.16: District AYP Passage on Turnover of Superintendents With More than One Year of Tenure (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	-0.0537	-0.0560	-0.0480	-0.0785	-0.0488	-0.0862
	(0.0533)	(0.0669)	(0.0689)	(0.0821)	(0.0599)	(0.0780)
Observations	3580	909	3580	909	3069	769
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.17: Better State Designation on Superintendent Turnover (RDD)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
Better State Desig.	0.00733	0.0237	-0.00561	0.0455	0.0173	0.0724
	(0.0186)	(0.0332)	(0.0217)	(0.0484)	(0.0274)	(0.0465)
Observations	3395	1781	2405	1046	1417	920
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Local Sample	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2008-2012. We report heteroskedasticity robust standard errors clustered at district-year level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conditional Logit Models

Table A.18: District Performance Information on School Board Turnover (Conditional Logit)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.0208	0.0227	-0.0467	0.0738
	(0.0666)	(0.0833)	(0.0758)	(0.109)
Academic Emergency	-0.503	-0.737	-0.174	0.243
	(0.333)	(0.494)	(0.606)	(1.214)
Continuous Improvement	-0.114	0.175	0.273	0.616
	(0.192)	(0.284)	(0.258)	(0.450)
Effective	-0.362^{+}	0.135	0.0883	0.551
	(0.214)	(0.314)	(0.281)	(0.482)
Excellent	-0.295	0.210	0.226	0.731
	(0.234)	(0.338)	(0.296)	(0.503)
Excellent w/ Distinction	-0.413	-0.183	0.208	0.717
	(0.268)	(0.372)	(0.332)	(0.537)
Observations	7313	7313	5890	5890
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report standard errors in parentheses. Academic Watch is the omitted state designation.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.19: Performance Information on Superintendent Turnover (Conditional Logit)

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	0.110	0.159	0.0963	0.129	0.0100	-0.0589
	(0.104)	(0.141)	(0.106)	(0.142)	(0.114)	(0.159)
Academic Emergency	0.693	1.353	-0.236	-0.348	0.136	-0.782
	(0.769)	(1.225)	(0.592)	(0.878)	(0.583)	(1.225)
Continuous Improvement	0.234	-0.196	0.0979	-0.0321	0.658^{+}	0.276
	(0.394)	(0.613)	(0.311)	(0.508)	(0.336)	(0.591)
Effective	0.183	-0.143	0.227	0.394	0.542	0.0754
	(0.423)	(0.652)	(0.343)	(0.553)	(0.371)	(0.637)
Excellent	0.272	0.241	0.213	0.462	0.673	0.125
	(0.445)	(0.675)	(0.376)	(0.589)	(0.411)	(0.687)
Excellent w/ Distinction	0.269	0.299	0.249	0.270	0.831	0.00101
	(0.488)	(0.725)	(0.449)	(0.685)	(0.510)	(0.871)
Observations	3640	3640	3640	3640	3024	3024
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. We report standard errors in parentheses. Academic Watch is the omitted state designation.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Districts With Consistent AYP and State Designations

Table A.20: District Performance Information on School Board Turnover (FE) No Conflicting Signals

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	-0.109	-0.0554	0.0246	0.0624
	(0.0788)	(0.126)	(0.0297)	(0.0475)
Academic Emergency	-0.103	-0.115	-0.188	-0.317
	(0.0773)	(0.147)	(0.152)	(0.258)
Continuous Improvement	-0.0377	0.0510	0.118	0.109
	(0.0529)	(0.0831)	(0.0714)	(0.0980)
Effective	0.0212	0.0975	0.120	0.120
	(0.0498)	(0.0707)	(0.0794)	(0.117)
Excellent	0.0546	0.134**	0.166+	0.131
	(0.0409)	(0.0510)	(0.0906)	(0.142)
Excellent w/ Distinction			0.168+	0.107
			(0.100)	(0.154)
Observations	4465	4465	3500	3500
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. We report heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.21: District Performance Information on School Board Incumbent Vote Share (FE) No Conflicting Signals

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI
District AYP Met	0.00535	0.115	0.0304	-0.0326
	(0.0597)	(0.116)	(0.0233)	(0.0416)
Academic Emergency	-0.110**	-0.0257	-0.0561	-0.106
	(0.0380)	(0.100)	(0.0462)	(0.0974)
Continuous Improvement	0.0757	0.168+	0.0283	0.0324
	(0.0470)	(0.0914)	(0.0518)	(0.0648)
Effective	0.0385	-0.00550	0.00252	0.0468
	(0.0354)	(0.0615)	(0.0658)	(0.135)
Excellent	-0.000835	-0.0265	-0.0401	0.104
	(0.0244)	(0.0402)	(0.0716)	(0.152)
Excellent w/ Distinction			-0.0313	0.147
			(0.0798)	(0.150)
Observations	925	925	776	776
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes

The unit of analysis in the models is individual incumbent school board members with performance information measured at the district level. The analysis includes years 2003-2012. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.22: Performance Information on Superintendent Turnover (FE) No Conflicting Signals

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	0.0464	0.125	0.0297	0.0692	0.00868	-0.0220
	(0.0867)	(0.133)	(0.0423)	(0.0731)	(0.0253)	(0.0368)
Academic Emergency	0.172	0.255	-0.0409	-0.150	0.00761	-0.00139
	(0.157)	(0.199)	(0.0983)	(0.131)	(0.111)	(0.151)
Continuous Improvement	-0.00683	0.0225	-0.0289	-0.0232	0.0646	-0.0336
	(0.0633)	(0.0843)	(0.0523)	(0.0801)	(0.0609)	(0.0907)
Effective	-0.0471	-0.0523	-0.0666	-0.0557	0.102	-0.00333
	(0.0436)	(0.0625)	(0.0605)	(0.0948)	(0.0688)	(0.113)
Excellent	0.0118	0.0514	-0.0538	-0.0492	0.116	0.0139
	(0.0325)	(0.0472)	(0.0683)	(0.107)	(0.0789)	(0.126)
Excellent w/ Distinction			-0.0615	-0.0834	0.133	0.0298
			(0.0777)	(0.116)	(0.0889)	(0.136)
Observations	2676	2676	2676	2676	2293	2293
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.23: Performance Information on Turnover of Superintendents With More than One Year of Tenure (FE) No Conflicting Signals

	(1) Current PI	(2) Current PI	(3) Lag PI	(4) Lag PI	(5) 2-Lag PI	(6) 2-Lag PI
District AYP Met	0.00343	0.0684	0.0236	0.0422	0.00377	-0.00825
	(0.0614)	(0.114)	(0.0329)	(0.0620)	(0.0198)	(0.0273)
Academic Emergency	0.239	0.329	-0.0918	-0.239*	0.0464	-0.0163
	(0.150)	(0.217)	(0.0591)	(0.115)	(0.0961)	(0.159)
Continuous Improvement	-0.0283	-0.0108	0.0302	0.00797	0.0490	0.0113
	(0.0400)	(0.0703)	(0.0320)	(0.0558)	(0.0438)	(0.0601)
Effective	-0.0226	-0.0511	-0.00388	-0.0172	0.0556	0.0132
	(0.0354)	(0.0493)	(0.0378)	(0.0637)	(0.0501)	(0.0749)
Excellent	-0.00522	-0.0160	-0.00129	-0.0132	0.0342	-0.00247
	(0.0261)	(0.0364)	(0.0459)	(0.0752)	(0.0585)	(0.0900)
Excellent w/ Distinction			-0.00927	-0.0285	0.0389	-0.0159
			(0.0533)	(0.0800)	(0.0669)	(0.0974)
Observations	2390	2390	2390	2390	2047	2047
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Time Trend	No	Yes	No	Yes	No	Yes

The unit of analysis in the models is individual superintendents with performance information measured at the district level. The analysis includes years 2004-2011. The Excellent with Distinction designation is used only starting in 2008. Heteroskedasticity robust standard errors clustered at district level in parentheses. Academic Watch is the omitted state designation.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$