NATIONAL
CENTER *for* ANALYSIS *of* LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

AIR

Labor Market Frictions and Production Efficiency in Public Schools

Dongwoo Kim
Cory Koedel
Shawn Ni
Michael Podgursky

WORKING PAPER 163 • December 2016

Labor Market Frictions and Production Efficiency in Public Schools

Dongwoo Kim
University of Missouri, Columbia

Cory Koedel
University of Missouri, Columbia

Shawn Ni
University of Missouri, Columbia

Michael Podgursky
University of Missouri, Columbia

Contents

Acknowledgements

They gratefully acknowledge research support from the Laura and John Arnold Foundation and the National Center for Analysis of Longitudinal Data in Education Research (CALDER) funded through grant #R305A060018 to American Institutes for Research from the Institute of Education Sciences, U.S. Department of Education.

The authors thank Maurice Li for research assistance and Dan Goldhaber, Peter Mueser and Eric Parsons for useful comments.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

Labor Market Frictions and Production Efficiency in Public Schools
Dongwoo Kim, Cory Koedel, Shawn Ni, Michael Podgursky
CALDER Working Paper No. 163
December 2016

Abstract

State-specific licensing policies and pension plans create mobility costs for educators who cross state lines. We empirically test whether these costs affect production in schools – a hypothesis that follows directly from economic theory on labor frictions – using geo-coded data from the lower-48 states. We find that achievement is lower in mathematics, and to a lesser extent in reading, at schools that are highly-exposed to state boundaries. A detailed investigation of the selection of schools into boundary regions yields no indication of systematic differences between boundary and non-boundary schools along other measured dimensions. Moreover, we show that cross-district labor frictions do not explain state boundary effects. Our findings are consistent with the hypothesis that cross-state mobility costs induce frictions in educator labor markets that lower student achievement.

1. Introduction

Several features of the labor market for public educators in the United States create mobility frictions. Within states, cross-district mobility can be hampered by the limited transferability of experience, which influences teacher placement on the salary schedule and other seniority-based benefits (e.g., preferences for open positions). Across states, teachers are subject to additional mobility costs owing to imperfect licensing reciprocity (Coggshall and Sexton, 2008; Goldhaber et al., 2015; Kleiner, 2015; Sass, 2015) and non-portable pension benefits (Costrell and Podgursky, 2010; Goldhaber et al., 2015; Koedel et al., 2012).[1] The research literature on educator mobility across state lines is thin, but what evidence is available is consistent with the additional costs of cross-state mobility impeding teacher movement. For example, a study of the Oregon/Washington border by Goldhaber et al. (2015) finds that cross-state teacher mobility is substantially lower than within-state mobility near the state line. Podgursky et al. (2016) document that cross-state teacher moves are rare in a study of three contiguous Midwestern states.

The additional mobility costs associated with crossing state boundaries for educators motivates the question of whether these costs introduce labor frictions that affect production. A number of studies of restricted labor mobility in other sectors point toward frictions lowering output (Botero et al., 2004; Caballero et al., 2013; Helpman and Itskhoki, 2010; Lafontaine and Sivadasan, 2009; Mitra and Ranjan, 2010). Moreover, in the education context specifically, Jackson (2013) shows that teacher-school match quality is an important determinant of teacher effectiveness, which implies that labor frictions that prevent some matches from occurring will be costly. We use geocoded

[1] It is also sometimes the case that an educator's seniority and tenure status will not carry over across a state line, in excess of any within-state mobility penalties along these lines. Goldhaber et al. (2015) document that this is true in Washington. However, in most cases seniority and tenure are determined at the district level, in which case a state change and district change would have similar effects.

information on schools in the United States merged with data on school-average achievement to empirically test whether exposure to state boundaries reduces output in schools.

We find highly-localized, robust negative effects of exposure to a state boundary on grade-8 student achievement in mathematics. Specifically, achievement at schools where a large share of the local-area workforce is on the other side of a state border is 0.09 *school-level* standard deviations lower on average than achievement at otherwise similar schools where none of the local-area workforce is outside of the state. We also estimate negative boundary effects on reading test scores, but these effects are about half as large as in math and are not as robust.

The key threat to identification in our study is that schools near state lines may differ systematically in other ways from schools that are farther away. We examine this possibility extensively using rich data from the National Center for Education Statistics (NCES) and the U.S. Census about schools and their local communities. There is no evidence that schools near state boundaries differ from other schools along measured dimensions within states. We also test whether our findings are driven by the presence of district boundaries – which coincide with state lines – independent of state-level factors. Although our models suggest that there may be costly frictions associated with district lines as well, district frictions cannot explain the state-boundary effects. Our results are consistent with cross-state mobility penalties for public educators inducing frictions in the labor market that lower student achievement.

2. Background

In this section we briefly discuss state-specific licensing and pension policies that impose costs on educator mobility across state lines.

2.1 Teacher Licensing

Teacher licensing requirements are set by state policy and typically specify that teachers attain a particular education level (e.g., a bachelor's degree), some form of state-approved preparatory experience, and/or pass one or more state certification tests (Sass, 2015). Although in many states an unlicensed teacher can teach under a temporary license for a short time, temporary licenses are usually not renewable and individuals who plan to have a career in teaching must obtain a state-specific license. There is significant variation in licensing tiers and types among states. Twelve states have just one licensing tier, nineteen states and the District of Columbia have two tiers, and nineteen states have three tiers. The tiers have different labels and requirements in different states. For instance, some states require additional coursework and a certain number of years of experience to move from a level-1 to level-2 license (e.g., Massachusetts) while others require a performance-based assessment (e.g., Washington). Differences in licensing policies across states tend to be largest for special education teachers, early childhood teachers, and middle school teachers (Coggshall and Sexton, 2008).

The Interstate Agreement (IA), created by the National Association of State Directors of Teacher Education and Certification (NASDTEC), reflects the policy concern that state licensing requirements restrict educator labor flows. The IA includes individual agreements between most US states that outline the process for obtaining a license for transfers (the IA covers 48 of the 50 states, plus the District of Columbia). Although the goal of the IA is to reduce licensing barriers to mobility, it does not offer full reciprocity; moving across state lines still requires additional steps to obtain a license in the new state (Coggshall and Sexton, 2008). Our review of the IA suggests that for many states the required steps are substantial (e.g., taking specific tests, completing new coursework, etc.). There are also typically more requirements for less-experienced teachers, which is important because labor mobility is higher among younger workers all else equal (Farber, 1999). Moreover, even in cases where reciprocity is complete or nearly so, we note that the general complexity of state teacher licensing rules can obscure this fact. Goldhaber et al. (2015) provide an example of a license that is fully reciprocal between Oregon and Washington, but for which reciprocity is not readily evident to a potential transfer. Numerous examples of complicated and unclear reciprocity conditions can be found in the IA.[2]

Recent work by DePasquale and Stange (2016) finds that a reduction in licensing barriers for nurses brought on by the Nurse Licensure Compact (NLC) did not increase cross-state labor mobility. One could interpret their findings as indirect evidence that licensing barriers are unimportant. However, like the IA, the NLC does not offer full licensing reciprocity and the literature is not clear on what aspects of imperfect licensing reciprocity drive behavior, making such inference somewhat speculative. It is also the case that DePasquale and Stange (2016) cannot rule out fairly large mobility effects of the NLC relative to the baseline mobility rate in some specifications.

2.2 Teacher Pensions

State-specific pension coverage is another source of cross-state mobility costs for public educators. Most teachers are enrolled in state defined-benefit (DB) pension plans, which are characterized by highly-backloaded wealth accrual (Koedel and Podgursky, 2016). The wealth-accrual backloading can result in severe financial penalties for teachers who switch plans.

There are two primary channels by which teachers' retirement plans penalize mobility. First, key plan benchmarks – vesting and retirement eligibility – depend on in-plan service years. Vesting rules typically require teachers to work 5 to 10 years in the same system in order to be eligible for a pension; if a teacher leaves prior to vesting she loses all employer contributions to the pension plan on her behalf (Backes et al., 2016). Retirement eligibility also depends on in-system service and individuals who split time in more than one plan usually must work longer to become eligible to collect a pension (Costrell and Podgursky, 2010). The other way that DB plans penalize mobility is through their calculations of the final average salary (FAS), which is used to determine the final pension

[2] See here for the IA: http://www.nasdtec.net/?page=interstate. Curran, Abrahams and Clarke (2001) discuss limitations of the IA with respect to its complexity and lack of symmetry between states.

payment a teacher receives. FAS is typically calculated as the average of the highest few years of earnings and is frozen at the time of exit. Thus, it does not account for inflation or life-cycle pay increases and this penalizes teachers who switch plans mid-career.

The precise costs facing a mobile teacher depend on the timing of the move and the details of the two plans, but Goldhaber et al. (2015) and Costrell and Podgursky (2010) document that cross-state mobility costs can routinely be upward of $100,000 in present value. Mobility costs are highest as teachers approach retirement but impact teachers throughout the experience distribution. Koedel et al. (2012) additionally provide evidence on pension mobility costs for school principals, which are even higher than for teachers owing to their higher late-career salaries. The high costs faced by school principals are notable in light of emerging evidence on the important role that principals play in educational production (Branch, Hanushek and Rivkin, 2012).

3. Empirical Strategy

We estimate the effect of state-boundary exposure on student achievement using a linear regression model of the following form:

$$Y_{ij} = \delta_0 + \mathbf{X_{ij}}\boldsymbol{\delta_1} + \mathbf{R_{ij}}\boldsymbol{\delta_2} + \gamma_j + \varepsilon_{ij} \qquad (1)$$

In Equation (1), Y_{ij} is average achievement on the state standardized test for school i in state j, normalized by state and grade. We estimate separate models for math and reading achievement. The vector $\mathbf{X_{ij}}$ includes rich information about schools and their local communities taken from the National Center for Education Statistics (NCES) and the U.S. Census. The full list of variables is shown in Table 1 and includes school- and district-average student demographics and socioeconomic-status measures, school and district enrollment, and district per-pupil revenue; for the local area, we include measures of population density, urbanicity, median household income and education levels. $\mathbf{R_{ij}}$ is a vector of exposure measures to the state boundary, for which we consider several different constructs as described in the next section. γ_j is a state fixed effect. ε_{ij} is the error term, which we cluster at the state level.[3]

The model in Equation (1) will yield unbiased boundary-effect estimates ($\boldsymbol{\delta_2}$) if boundary exposure is independent of the error term conditional on observed covariates in the *X*-vector; i.e., selection-on-observables. Although it is not possible to test directly for unobserved selection into boundary regions, in Section 5 we show that there is no evidence of selection along any of the observed dimensions measured by the rich NCES and the U.S. Census datasets. The results from our analysis of observed selection imply that unobserved selection is also likely to be of limited practical importance (Altonji, Elder and Taber, 2005).

We use grade-8 test scores as the primary outcomes of interest. Students typically do not stay in the same school through grade-8, but given the local nature of the provision of public schooling in the U.S., boundary closeness in middle school is indicative of boundary closeness in earlier grades as well. Thus, the boundary effects we estimate are best viewed as cumulative effects of repeated exposure for students. Grade-8 is the highest consistently-tested grade in U.S. public schools, making it the grade in which we are most likely to see boundary effects that have accumulated over time in available testing data. We also estimate the effect of boundary closeness on test scores in earlier grades. If the labor-frictions mechanism is correct and boundary effects accumulate then we should estimate smaller effects in earlier grades, which is indeed what we find.

4. Data

4.1 Defining Schools' Local Labor Markets

We geocode the locations of schools with respect to state boundaries and other schools in their local geographic areas in the lower 48 states. Our preferred measure of how much each school's

[3] In Equation (1) and all subsequent equations, variables are in row vectors and parameters are in column vectors.

local-area labor market is exposed to a state boundary is constructed in the following steps. First, we draw a circle around each school with a 10-mile radius. Next, we identify the total number of full-time equivalent (FTE) teachers at other schools within each reference school's circle using the Common Core of Data (CCD) maintained by the NCES. We then determine the share of FTE teachers who work in another state within the 10-mile radius. Figure 1 provides an illustrative example. In the figure, School A is the reference school and Schools B, C, D, and E are within its local-area circle. School E is on the other side of a state line. In this case, our measure of the intensity with which School A's local-area labor market is affected by the boundary is

$$F_E / (F_B + F_C + F_D + F_E) \qquad (2)$$

where F_X is the number of FTE teachers at School X as reported in the CCD. If School A were far from a state boundary, the value of this measure would be zero because all nearby schools would be in the same state.

The rationale underlying our measure is that for a given geographic labor market, frictions brought on by a state boundary will shrink the effective labor pool. Equation (2) is a measure of the intensity of boundary exposure motivated by the large research literature examining how labor frictions affect firm behavior and productivity (Botero et al., 2004; Caballero et al., 2013; Helpman and Itskhoki, 2010; Lafontaine and Sivadasan, 2009; Mitra and Ranjan, 2010).[4] These previous studies

[4] Moreover, because schools offer similar salaries, nonpecuniary job features are particularly important in the educator labor market. Greenberg and McCall (1974) note that this makes internal mobility for teachers more valuable. New teachers may be more likely to start in less desirable schools with plans to move to more desirable schools as they become more experienced. By making some local-area mobility options more costly, state boundaries will make positions at nearby schools less desirable, which in turn will lower the quality of the applicant pool for boundary schools relative to non-boundary schools, ceteris paribus.

motivate the question we aim to answer: Does exposure to frictions from state boundaries affect schooling output?

Of course, there are other reasonable ways to construct measures of boundary exposure. Correspondingly, we consider the robustness of our findings to a number of alternative measures. We replace our FTE-based measures with measures based on school-level student enrollment. We also use measures that are restricted to include only other schools in the reference school's local area with overlapping grades, and schools with similar student populations (measured by the share eligible for free/reduced-price lunch). Our circles with 10-mile radii are supported by a small literature on teacher commuting showing that most teachers do not commute far, but we also construct alternative measures of boundary exposure using circles with 20-mile radii and incorporate them into the analysis as well.[5] In addition, we construct models that use count-based measures instead of ratio-based measures of boundary exposure, models that aggregate schools up to the district level to examine whether boundary exposure at the district level influences achievement, and perform several other tests as detailed below. Overall, our findings are robust to a variety of ways of measuring and modeling the extent to which a school's local-area labor market is exposed to a state boundary.

Using our preferred measures based on a 10-mile radius, roughly 12 percent of schools in the U.S. have non-zero boundary exposure (i.e., have at least one school in their 10-mile local area on the other side of a state line). Approximately 5 percent of schools have 25 percent or more of their total local-area FTE on the other side of a state line, and throughout our analysis we identify this group as "intensely affected" by a boundary. The distribution of the out-of-state FTE percentage, constructed using Equation (2), is shown in Appendix Figure A1 for schools in our analytic sample. Unfortunately,

[5] Evidence on average commuting distances for public school teachers in the U.S. is rare, but what evidence is available suggests teachers have strong preferences for short commutes. For example, in their analysis of a large urban district, Miller, Murnane and Willet (2008) find that the average teacher commutes just 7 miles to work. Using applicant data from Chicago, Engel, Jacob and Curran (2014) find that teachers are 40 percent less likely to apply to an opening at a school that is just over three miles further from their homes (for related evidence also see Cannata, 2010).

the measure does not afford much flexibility in how we define "intensely affected" schools. For example, if we change the threshold for FTE on the other side of a state line from 25 to 50 percent, the share of schools that satisfy the criterion falls by more than half, from roughly 5 to 2 percent. To illustrate why the sample size declines so quickly as we increase the threshold, consider a stylized example of a school near a single, straight-line state boundary (like in Figure 1). Imagine that the school is surrounded by equal-sized schools that are distributed in a geographically uniform manner across the local area. Because the circle we draw around the school is centered on itself, and the school is in its own state (obviously), the out-of-state area covered by the circle must be less than the in-state area, and thus in expectation the out-of-state FTE percentage will be smaller than the in-state-FTE percentage. As a practical matter, this results in the number of schools categorized as "intensely affected" by a boundary declining rapidly as we increase the FTE threshold, as illustrated in the appendix. Consistent with this measurement issue, in unreported results we find that further subdividing the group of intensely affected schools does not yield additional insights because statistical power is significantly reduced.

A last issue that merits brief mention involves sparsely populated areas where schools operate but there are no other schools within 10 miles. There are few such schools – approximately 1.9 percent of the schools in our sample – and we code them as being unaffected by state boundaries. That is, we code their local area out-of-state FTE share as zero (although the ratio in Equation (2) would technically be undefined). This coding is to reflect the fact that even though these schools surely face thin educator markets, the thinness is not driven by state boundaries, which we leverage for identification. The theoretical implications of local-area labor market thinness for these schools are likely the same as for any other school – including schools near state boundaries – but the concern is that these schools also differ in other ways from schools in thicker markets. Thus, we do not leverage them to identify the parameters of interest in our study, which we restrict to be identified off of state-boundary induced local-area segmentation. We note that our findings are also robust to simply dropping schools in sparsely populated areas from the sample, which we show in Appendix Table A6.

4.2 Achievement

We estimate the effects of boundary proximity on school-average grade-8 standardized test scores in math and reading from the 2012-2013 school year. We normalize scores within state-grade-subject cells. Because the "treatment" in our case is time invariant and school-average test scores are highly serially correlated, adding additional years of outcome data is of little practical value in our application (Bertrand, Duflo and Mullainathan, 2004).[6] We collected test score data from state departments of education online and via direct correspondence. Some states did not have the data, were unwilling to process our request, or we were unable to use the data. Ultimately, we use standardized test score data from 33 of the lower-48 states in our primary models.[7]

We also estimate the effect of boundary closeness on school proficiency rates. We normalize school proficiency rates within state-grade-subject cells as well. A benefit of using proficiency rates is that they are more commonly available from state education agencies and allow us to extend our analytic sample to include 43 of the lower-48 states (see Appendix Figure A2 and Appendix Table A2 for more information about our sample coverage using the proficiency rate data). Our findings are substantively similar using standardized test scores and proficiency rates. That said, while the use of proficiency rates allows us to increase the coverage of our analytic sample, there are well-documented

[6] That said, we collected data from schools in a subsample of states during the 2013-2014 school year to confirm that, however unlikely, our findings are not driven by a peculiarity in the 2012-2013 data. As expected, the 2013-2014 results look very similar to what we find using the 2012-2013 data.

[7] We collected data from 35 states, but we cannot use data from (a) Missouri, (b) California for grade-8 math, and (c) Nebraska for grade-8 reading. Missouri is in the unique situation of having more than one pension plan within the same state without reciprocity (Koedel et al., 2012). Our geocoded data cannot capture the pension boundaries within the state, and for this reason we exclude Missouri. In California, the grade-8 test data in math are not as useful as in other states because of the strong push in California to have grade-8 students take algebra-I, and thus the algebra-I test in place of the typical grade-8 standardized exam (Domina et al., 2014). There is significant variation across California schools in the proportion of students taking the algebra test, and the overall rate of standardized test taking is much lower than in other states. In a robustness test shown below, we are able to bring California data into our analysis by estimating boundary effects on grade-7 math scores. Finally, we do not have reading test data from Nebraska because they are not accessible online and the Nebraska Department of Education did not respond to our data request.

measurement issues associated with proficiency rates and for this reason we do not emphasize these results too strongly (Bandeira de Mello, 2011; Bandeira de Mello et al., 2015; Ho, 2008).

Figure 2 shows the 33 states included in our primary analytic sample with grade-8 standardized test data in mathematics. Table 1 compares the schools in our 33-state sample to the full sample of schools in the lower-48 states.[8] The school- and district-level data are from the 2012-2013 CCD and zip-code level data are from the 2013 American Community Survey (ACS) 5-year estimates. While there are some differences between our sample and the national sample, they are generally similar. The bottom rows of the table show that the share of schools nationally for which a state boundary bisects the 10-mile circle is very similar to the share of schools in our sample, as is the share of schools for which 25 percent or more of the local-area FTE is on the other side of a state line. Note that while "intensely affected" boundary schools make up just a small fraction of our sample (≈5 percent), they account for many students. Just based on middle-school students, enrollment in these schools nationally during the 2012-2013 school year was approximately 670,000 students, which is roughly equivalent to total middle school enrollment in the three largest school districts in the country combined (New York, Los Angeles, and Chicago).

5. Results

5.1 Selection into Boundary Regions

We begin by examining selection of schools into boundary regions; i.e., whether schools with more intense exposure to a state boundary differ from schools with less (or no) boundary exposure along observed dimensions. Endogenous selection into boundary regions, or any geographic region for that matter, is likely less of an issue for schools than for other entities – e.g., private firms – *a priori*

[8] Per above (footnote 7), for reading scores we include California in the analytic sample and remove Nebraska. We do not report separate sample characteristics for the math and reading samples because they overlap entirely except for these two states and thus are very similar. Appendix Table A1 provides additional details about the construction of our analytic sample.

because schools must cover all geographic areas. Nonetheless, it may still be that schools near state boundaries differ from other schools. We examine this possibility in two related ways.

First we use predicted test scores based on observed school, district, and local-area characteristics as summary measures of baseline characteristics to compare boundary and non-boundary schools. We start by estimating the following supplementary regression:

$$Y_{ij} = \tau_0 + \mathbf{X_{ij}}\boldsymbol{\tau_1} + \psi_j + e_{ij} \tag{3}$$

The covariates in Equation (3) are as defined in equations (1) and (2) and are strong predictors of test scores – for grade-8 math and reading scores, the R-squared values from Equation (3) are 0.48 and 0.58, respectively. We use the output from Equation (3) to construct a predicted test score for each school based on observable characteristics, $\hat{Y}_{ij} = \hat{\tau}_0 + \mathbf{X_{ij}}\hat{\boldsymbol{\tau}}_{\mathbf{1}} + \hat{\psi}_j$. The gaps in predicted test scores between intensely-affected boundary schools and other schools are very small: 0.011 and 0.023 school-level standard deviations in math and reading, respectively. Both gaps are statistically insignificant and nominally favor intensely affected boundary schools.

We provide an expanded analysis of selection using variants of the following regression model:

$$\tilde{R}_{ij} = \lambda_0 + \mathbf{X_{ij}}\boldsymbol{\lambda_1} + \rho_j + u_{ij} \tag{4}$$

In Equation (4), $\tilde{R}_{ij}$ is a measure of boundary closeness for school *i* in state *j*, the vector $\mathbf{X_{ij}}$ includes the same school-level covariates used in Equation (1), ρ_j is a state fixed effect and u_{ij} is the error term. Non-zero entries in the parameter vector $\boldsymbol{\lambda_1}$ are indicative of selection into boundary regions, as measured by $\tilde{R}_{ij}$, along observed dimensions within states.

We estimate Equation (4) with and without state fixed effects, and defining $\tilde{R}_{ij}$ and the analytic sample in several ways. From the discussion in Section 4.1, our primary models (Equation 1) divide schools into three groups: (1) those with 25-percent or more of the local geographic labor market in another state, (2) those with more than zero but less than 25 percent of the local geographic labor market in another state, and (3) those with none of the local geographic labor market in another state. In Table 2 we report results from Equation (4) where we define $\tilde{R}_{ij}$ to inform selection based on these groupings, along with results from a model where we code $\tilde{R}_{ij}$ as a simple linear variable indicating the share of local-area FTE on the other side of a state line. The out-of-state FTE share is not well-captured by the linear variable (see Appendix Figure A1), but the linear measure affords an alternative way of looking for evidence of selection. We also show several other versions of the selection model in Appendix Table A3, which corroborate the findings in Table 2.

The top rows of Table 2 show full output for each model. At the bottom of the table, we report p-values for the likelihood of observing the number of unbalanced covariates indicated in the model by chance, at the 10 percent level, in the state-fixed-effects specifications. The p-values are generated using randomized inference as in Cullen, Jacob and Levitt (2006) and Fitzpatrick, Grissmer and Hastedt (2011) and account for the covariance structure of the data.[9] We show results in Table 2 using our primary analytic sample for math, but in an omitted analysis we confirm that our findings are similar if we use the analytic sample for reading.[10]

In the first two sets of results in Table 2 (columns 1-4), we define $\tilde{R}_{ij}$ as an indicator variable equal to one if school *i* is an "intensely affected" school per the above definition, and zero otherwise.

[9] To obtain the randomized-inference p-values, we start by splitting the analytic dataset vertically, separately blocking off the covariates (independent variables) and the measures of boundary closeness. The vertical blocking maintains the covariance structure between the variables in the *X*-vector, which is important because the covariance structure will influence the probability of observing any given number of statistically significant relationships with the real data. We randomly sort the block of covariates, then re-connect it to the block of boundary-closeness measures, which effectively assigns each school a random boundary-closeness measure. We then run the model in Equation (3) and store the number of unbalanced covariates obtained under random assignment. We repeat this procedure 3,000 times to construct empirical distributions of covariate imbalance, from which the p-values are obtained.

[10] Per footnote 7, the only difference between the math and reading samples in grade-8 is that the math sample excludes California and includes Nebraska, and vice versa for the reading sample. Unreported sensitivity analysis, corroborated by our models of grade-7 test scores below, confirm that this single-state substitution has no qualitative bearing on our findings.

In columns 1-2 we include all other schools (i.e., the remaining ≈95-percent of our sample) in the model and assign them values of zero for $\tilde{R}_{ij}$. This comparison treats moderately affected schools (those with more than zero but less than 25 percent of the local-area labor market on the other side of a state line) like any other control school. In columns 3-4 we drop moderately affected schools from the sample entirely, which results in our comparing intensely affected schools to the remaining 88 percent of schools for which no other school within the 10-mile circle is on the other side of a state line. The models in columns 5-6 include the full sample of schools but use the linear out-of-state FTE share as the dependent variable.

While there is some evidence in Table 2 of imbalance between boundary and non-boundary schools owing to cross-state differences (i.e., in columns 1, 3, and 5), the models with state fixed effects provide no evidence of selection into boundary regions. The p-values reported at the bottom of the table are well above conventional levels of significance in all cases, ranging from 0.26-0.69. These balancing results are achieved despite the fact that in most cases our standard errors are not large, and decline for many covariates when we move to the state-fixed-effects specification. Based on these results, we conclude that there is no evidence of selection into boundary regions along the measured dimensions of our data, which we again note are quite rich.

5.2 Primary Results for Grade-8 Achievement

Table 3 shows the effects on math and reading achievement of exposure to a state boundary as estimated by Equation (1). Coefficients for the non-boundary covariates are suppressed in Table 3 (and subsequent tables) but can be found in Appendix A. Again, our standard errors are clustered at the state level. Focusing first on the grade-8 math model, we find that intense exposure to a state boundary lowers student achievement by 0.094 school-level standard deviations. In reading, test scores in intense-exposure schools are 0.054 standard deviations lower than in non-boundary schools and the difference is marginally significant. Consistent with the similarity of schools that differ by boundary exposure as documented in the preceding section, in Appendix Table A5 we show that the findings in Table 3 are not sensitive to which components of the X-vector are included in the model.

The effect sizes in Table 3 (and subsequent tables) are reported in standard deviations of the distribution of school-average achievement, which are akin to what one might estimate in a study of firm-level productivity. In education research, effect sizes are typically reported in student-level standard deviation units. Bhatt and Koedel (2012) find that a scaling factor of roughly one-third translates effect sizes in the school-level distribution to the student-level distribution. In our application, this would imply that the 0.094 effect size in the distribution of school-average math scores would translate to a 0.031 effect size in student-level standard deviations.[11]

The control group in Table 3 consists of schools for which the entire local-area labor market is within a state. The results suggest that schools where a small fraction of the local-area labor market is on the other side of a boundary – above zero but less than 25 percent of surrounding FTE – are not affected in the same way as intensely affected schools. Although we cannot rule out modest negative effects for these schools given the sizes of our standard errors, the weaker findings persist through many robustness and sensitivity analyses below, which implies that the boundary effect is highly localized.

In Table 4 we report analogous results where we use school proficiency rates on state tests in place of the standardized test scores. Our proficiency rate measures indicate the share of grade-8 students in the school rated as proficient or above on the state assessment and are standardized within states and subjects. Proficiency rate data are available at the school level in 43 of the lower 48 states, which affords a significant expansion of our sampling frame. The appendix provides additional details

[11] Burgess, Wilson and Worth (2013) use a similar scaling factor to move between school- and student-level test-score distributions in a different context. Note that the standard deviations of test scores at the school and student levels include variance due to measurement error and the measurement error variance will be larger in student-level scores. Thus, effect sizes in the true distributions of achievement are larger (Boyd et al., 2008).

about the sample expansion. Table 4 presents proficiency-rate models that restrict the sample to include only states from Table 3 (i.e., states for which we have school-level standardized test scores), as well as models that use the broader sample afforded by the proficiency data. The findings are generally consistent with the results in Table 3. They indicate an effect of boundary closeness on the math proficiency rate for intensely affected schools of -0.082 standard deviations, and a statistically insignificant effect of -0.033 standard deviations in reading. Although the positive estimate for schools with more than zero but less than 25 percent of local FTE on the other side of a state line in Table 4 for the extended sample in reading is peculiar, this result is not replicated anywhere else in our analysis and thus we do not put much weight on its significance (in particular, see Tables 3, 5, 8 and 9, along with the battery of tests in Appendix B).

6. Robustness and Extensions

6.1 Robustness and Sensitivity Analysis

As noted previously, we examine the robustness of our findings to a variety of ways of measuring and modeling boundary exposure and comparing schools. We relegate most of the robustness analysis to Appendix B, where we consider: (a) measuring boundary exposure in terms of local-area school enrollment instead of local-area FTE teachers, (b) restricting the exposure measures to include only schools with overlapping gradespans, (c) restricting the exposure measures to include only schools with similar student-body compositions as measured by the share of free/reduced-price lunch eligible students, (d) using exposure measures that include the school's own local-area FTE in the denominator of Equation (2), (e) the use of an alternative, more-differentiated control group, (f) models that use linear and quadratic measures of boundary exposure based on the out-of-state FTE share, and (g) models that use the simple distance to a state boundary (linear and quadratic terms). Summarizing the laundry-list of results in the appendix, our findings are qualitatively robust to the various modeling and measurement modifications.

In Table 5 we report on a sensitivity test related to the local intensity of boundary effects suggested by our estimates in Table 3. Specifically, we incorporate information from local-area labor markets defined by circles with 20-mile radii by estimating models of the following form:

$$Y_{ij} = \alpha_0 + \mathbf{X_{ij}}\boldsymbol{\alpha_1} + \mathbf{R_{ij}^{10}}\boldsymbol{\alpha_2} + \mathbf{R_{ij}^{20}}\boldsymbol{\alpha_3} + \theta_j + \xi_{ij} \qquad (5)$$

Equation (5) is the same as Equation (1) except that we include measures of boundary exposure using 10- and 20-mile radii simultaneously. The parameter vector $\boldsymbol{\alpha_3}$ indicates how having a higher percentage of local-area FTE in the 20-mile radius on the other side of a state line, conditional on the percentage of FTE outside of the state in the 10-mile radius, affects achievement.

The results in Table 5 reinforce the point from above that the boundary effects are concentrated. Conditional on how the local-area labor market is affected by a state boundary using the 10-mile radii, differences in how the market is split within the 20-mile circle has no discernable effect on achievement. Adding the 20-mile information increases the standard errors on our 10-mile measures due to the fact that both measures contain correlated information, but the point estimates for the 10-mile measures remain similar to what we show in Table 3.

In Table 6 we report estimates from alternative models that uses count-based measures of local-area FTE in place of the ratio-based measures we have used thus far. The count-based approach is similar to the previous approach except that we include the numerator and denominator of the ratio in Equation (2) as separate terms (as in Fitzpatrick and Lovenheim, 2014). The models take the following form:

$$Y_{ij} = \gamma_0 + \mathbf{X_{ij}}\boldsymbol{\gamma_1} + \mathbf{FTE_{ij}}\boldsymbol{\gamma_2} + \mathbf{FTE_{ij}^{OS}}\boldsymbol{\gamma_3} + \pi_j + e_{ij} \qquad (6)$$

The variable vectors $\mathbf{FTE_{ij}}$ and $\mathbf{FTE_{ij}^{OS}}$ in Equation (6) include linear and quadratic terms that measure the raw number of FTE within the 10-mile radius, and within the 10-mile radius and outside the state (OS), respectively. All other variables are the same as in Equation (1). The count-based model offers an alternative way to measure the impact of boundary closeness. A benefit of using this model is that the general relationship between having more FTE nearby and student achievement can be estimated at the same time as the boundary effect. A caveat is that $\boldsymbol{\gamma}_2$ may not be causal; it is identified using variation between incomparable schools that differ by labor-market thickness for reasons other than closeness to a state boundary. However, conditional on $\mathbf{FTE_{ij}}$, $\mathbf{FTE_{ij}^{OS}}$ is plausibly exogenous.[12]

The results in Table 6 are consistent with what we have shown thus far. The first two columns document the relationship between local-area FTE and student achievement, omitting information about the location of this FTE with respect to state boundaries (i.e., the models are estimated without $\mathbf{FTE_{ij}^{OS}}$). The relationship between local-area FTE and achievement is positive and weakly concave. When we add the state boundary information, the total FTE coefficients are similar and out-of-state FTE has a negative effect on achievement conditional on total FTE. At average in-state and out-of-state FTE values for intensely-affected boundary schools and control schools (control schools have zero out-of-state FTE), the coefficient estimates in Table 6 imply a test-score difference between school types of approximately -0.080 standard deviations, which is very close to the ratio-based estimate in Table 3.

6.2 Extensions

We extend our analysis in three ways. First, using our original boundary-exposure measures, we examine effect heterogeneity across local areas that differ by population density. Although we condition directly on urbanicity and population density in our main models, it may also be that there is heterogeneity in the importance of state boundaries depending on whether the school is in a high- or low-population-density area. To test for heterogeneous boundary effects we add interaction terms

[12] Another advantage of the count-based model is that it allows for what is perhaps a more intuitive, or at least more consistent, way of treating rural schools without any other schools in the 10-mile radius. As a practical matter this benefit is marginal because as noted above, there are few such schools and how we handle them in the previous models has no bearing on our findings (see Appendix Table A6).

between the boundary-exposure indicators and indicators that divide schools by population density. Specifically, we split schools into above- and below-median groups based on the population density of their zip-code areas.

Table 7 shows results from the interaction models, where we report the marginal effect of being in a lower-density area. We find no evidence to suggest that boundary effects are heterogeneous, but a caveat to this interpretation is that we are not powered well enough to examine the potential for heterogeneous effects in great detail. For example, we cannot rule out the possibility that boundary effects are particularly important in the most sparsely-populated areas because we do not have enough schools in such areas to obtain estimates with sufficient precision. Even with the broad cut we make to the data in Table 7, our standard errors on the heterogeneity parameters are large and cannot rule out substantial heterogeneity.[13]

The next extension we consider is to look at boundary effects in lower grades – grade-7, grade-5 and grade-3. There are two benefits of these estimates. First, we can expand our analytic sample for math in earlier grades to include California, which we dropped from our analysis of grade-8 due to the test-coverage issue discussed in Section 4.2. Second, and more importantly, the early-grade models provide indirect evidence about whether labor frictions are likely to drive our findings. Recall from above that labor frictions should have a cumulative impact given the local delivery of education services – i.e., attendance at a boundary school in grade-8 likely implies attendance at a boundary school in earlier grades as well. Because each year of exposure should influence total achievement if

[13] We also estimate models designed to examine heterogeneity in boundary effects owing to cross-state differences in teacher salaries (based on average salary data by state obtained from the NCES Digest of Education Statistics). The models test whether mobility frictions near boundaries can be offset or strengthened by differential salaries. We find suggestive evidence that boundary effects are more severe for schools in low-salary states that share boundaries with high-salary states. However, like with the density-based heterogeneity analysis, the estimated heterogeneity parameters are too imprecise to draw meaningful inference and thus we omit the results. We encounter similar statistical power problems in other attempts to examine boundary-effect heterogeneity.

labor frictions are responsible for our findings, it follows that boundary effects in lower grades should be smaller than in grade-8.[14]

Table 8 shows boundary effects in grade-7, grade-5, and grade-3. The grade-7 point estimates are similar but slightly smaller than the garde-8 estimates shown above. The estimates for grades 5 and 3 are even smaller, and statistically insignificant in both subjects.[15] The pattern of the estimates is consistent with the hypothesized cumulative nature of the effect of boundary exposure. In contrast, it is not consistent with a story that unobserved selection drives our findings, in which case there would be no reason to expect differences in the estimates by grade. We also note that our weaker results in reading throughout are consistent with a labor-frictions explanation – a large body of research shows that teachers and teacher-related interventions have smaller effects on reading achievement than math achievement (e.g., Hanushek and Rivkin, 2010; Lefgren and Sims, 2012; Taylor and Tyler, 2012).

Finally, we construct models to look for evidence of boundary effects in district-level test data. For each school district, we build aggregated boundary-exposure measures based on our school-level measures from Equation (2). These district-level measures capture the share of intensely or moderately affected boundary schools in each district. As an example, a district with 10 schools, one of which is intensely affected by a state boundary, would have an intense-exposure share of 0.10.

We use normalized estimates of district-average test scores as outcomes in the district-level models. The district test score estimates taken from the publicly-available Stanford Education Data Archive (SEDA; Reardon et al., 2016a) and are derived from information about student performance

---

[14] Another frictions-related factor that may contribute to smaller estimates in lower grades is that teaching positions may be easier to fill in elementary versus middle schools. To the extent that boundary effects are moderated by the underlying thickness of the labor market, the grade-8 effects will encapsulate the more pronounced effects of boundary closeness during the middle-school years.

[15] The structure of the schooling system is such that our elementary analysis includes many more schools (multiple elementary schools typically feed into a single middle school), but this does not improve precision because of the clustering structure of the data at the state level.

across all proficiency levels within states (Reardon et al., 2016b). A benefit of using the SEDA is that data are available for all lower-48 states (except California in grade-8 math for the same reason that we exclude California in our grade-8 math models), which affords another opportunity to look at an expanded sample (similarly to the proficiency-rate analysis).

Table 9 shows results from district-aggregated models that follow the format of our previous results. We show district-level estimates for the restricted sample of states for which we have school-level standardized test scores, and the expanded sample that includes all lower-47/48 states (again, California is excluded in the math model because SEDA data are unavailable). The results are consistent with our school-level findings and very similar using the restricted and full samples. In the math model with all 47 states, the estimate in Table 9 implies that going from a 0 to 1.0 share of intensely affected schools corresponds to a reduction in district test scores of 0.043 district-level standard deviations; the corresponding number in reading is 0.026 standard deviations. The math estimates are statistically significant at the 5 percent level in both samples. The reading estimates are statistically significant at the 5 percent level in the full sample and on the margin of statistical significance (p-value ≈ 0.14) in the restricted 33-state sample.

7. State Boundaries or District Boundaries?

Thus far we have established that schools with a large fraction of local-area FTE on the other side of a state line have lower achievement than otherwise similar schools where the labor market is not bisected by a state boundary. There is no indication that schools with more and less boundary exposure differ along other dimensions. The motivation of our study is to test for effects on schooling output that are predicted by economic theory if state-specific licensing and pension policies create labor frictions near state boundaries. However, district boundaries necessarily coincide with state lines and can induce their own frictions (e.g., due to imperfect mapping across salary schedules, general frictions associated with changing employers, etc.), making them a potentially confounding factor.

To test whether our findings are driven by frictions associated with district boundaries, we add control variables to our models that measure exposure to district boundaries directly. Specifically, we re-estimate Equation (1) with added controls that are analogous to our out-of-state FTE share variables, but capture the share of local-area FTE outside of *the district*.

$$Y_{ij} = \beta_0 + \mathbf{X_{ij}}\boldsymbol{\beta_1} + \mathbf{R_{ij}^{\mathit{OD}}}\boldsymbol{\beta_2} + \mathbf{R_{ij}^{\mathit{OS}}}\boldsymbol{\beta_3} + \varphi_j + \varepsilon_{ij} \tag{7}$$

In equation (7), the variables in $\mathbf{R_{ij}^{\mathit{OD}}}$ measure the out-of-district FTE share and the variables in $\mathbf{R_{ij}^{\mathit{OS}}}$ measure the out-of-state FTE share. All other variables are as defined above. An out-of-state school in the local area is necessarily out-of-district because no district spans state lines, but the reverse is not true. Thus, when both controls are included simultaneously, the coefficients on the out-of-district FTE variables $(\boldsymbol{\beta_2})$ are identified entirely from district boundaries within states. The coefficients on the out-of-state FTE variables $(\boldsymbol{\beta}_3)$ capture the influence of the out-of-state FTE share conditional on the out-of-district FTE effect.

The results are shown in Table 10. An important clarification for interpreting the results is that the omitted comparison group changes substantially in equation (7). Specifically, the omitted group now includes only schools that are not exposed to any FTE outside the *state or district* within 10 miles. Because most districts cover small geographic areas, this group is much smaller and more selected than in previous models. Specifically, just 15.8 percent of the schools in our sample have an out-of-district FTE share of zero, and more than two-thirds of schools have an out-of-district FTE share above 25 percent.[16]

[16] Clearly the distribution of out-of-district FTE is quite different than the distribution of out-of-state FTE (the latter distribution is documented in Appendix A). This suggests that a different way of codifying exposure to district boundaries may be warranted. Correspondingly, we have considered a variety of ways of controlling for exposure to district boundaries simultaneously with state boundaries and the qualitative implications are always similar to what we report in the main text. Given the similarity of results, we use an analogous coding scheme for out-of-state and out-of-district FTE shares for presentational convenience in Equation (7). An extended treatment of this issue is omitted from the manuscript for brevity but more information is available from the authors.

The most important takeaway from Table 10 is that district-boundary frictions do not drive our findings for state boundaries – the effect of intense exposure to a state boundary is essentially the same in Table 10 as it is in Table 3. The coefficients on the out-of-district FTE variables are negative, which implies that district boundary frictions may also lower achievement (these parameters are also identified much more precisely owing to the substantial within-state variation that informs the estimates). Combining the out-of-district and out-of-state FTE share coefficients would imply that schools with intense exposure to a state boundary have much lower achievement than schools within states that are exposed to neither state nor district boundaries – i.e., the implied effect is a relative reduction of test scores in math of -0.1447 (-0.0852 + (-0.0595)) school-level standard deviations, or roughly -0.05 student-level standard deviations.[17]

8. Conclusion

We study the effect on student achievement when a school's local-area labor market is bisected by a state boundary. We find robust and highly-localized negative effects of intense exposure to a state boundary on the order of 0.09 school-level standard deviations of grade-8 math test scores. In reading, we find smaller effects of about 0.05 school-level standard deviations that are only sometimes statistically significant. These estimates can be translated into student-level standard deviations by multiplying them by roughly one-third (Bhatt and Koedel, 2012; Burgess, Wilson and Worth, 2013). Although this is a small per-student effect, it is spread across a very large population: based on the Common Core of Data, we estimate that roughly 670,000 students are enrolled in middle schools nationally that would be coded as "intensely affected" by a state boundary in our study.

[17] We make this interpretation cautiously because unlike with the out-of-state FTE shares, the out-of-district FTE shares are correlated with other school characteristics within states and thus subject to greater concerns of selection bias. We omit results from selection equations for the out-of-district FTE share for brevity, but they are available from the authors.

Labor frictions at state boundaries are a plausible explanation for our findings. A large literature in economics documents the adverse effect of restricted labor flows on production (Botero et al., 2004; Caballero et al., 2013; Haltiwanger, Scarpetta and Schweiger, 2006; Helpman and Itskhoki, 2010; Lafontaine and Sivadasan, 2009; Mitra and Ranjan, 2010) and explicit state policies make it costly for educators to cross state lines. Our empirical results are consistent with what would be predicted by economic theory in this regard.[18] Moreover, one might also hypothesize that boundary effects on achievement will be more pronounced where labor markets are inherently thin, such as high school math and science teachers. Unfortunately, comprehensive testing data are not currently available on a national level to test for effects in higher grades where some types of labor may be particularly scarce.[19]

[18] In principle there is variation in the degree of difficulty in obtaining a license across any given state line, and in the pecuniary pension cost associated with the crossing, that could be exploited to gain further insights into mechanisms. However, it is unclear how well teachers understand these differences (e.g., see Chan and Stevens, 2008; Clark, Morrill and Allen, 2012; Goldhaber et al., 2015). The complexities of the rules are apparent in a perusal of the Interstate Agreement on licensing and the various pension plan reports. Moreover, even if teachers did understand the variation in costs, given the current size of our standard errors (affected primarily by the fundamental clustering structure at the state level) it is unlikely that sufficient statistical power is available to disentangle these competing mechanisms with reasonable precision in the current framework without imposing structural assumptions.

[19] Tests are administered to high school students but they vary within and across states in purpose, coverage (e.g., many tests in high school are not compulsory), and timing (e.g., see Parsons et al., 2015), which makes a national analysis using high school test data challenging.

References

Altonji, Joseph G., Todd E. Elder and Christopher R. Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113(1), 151–84.

Backes, Ben, Dan Goldhaber, Cyrus Grout, Cory Koedel, Shawn Ni, Michael Podgursky, P. Brett Xiang and Zeyu Xu. 2016. Benefit or Burden? On the Intergenerational Inequity of Teacher Pension Plans. *Educational Researcher* 45(6), 367-377.

Bandeira de Mello, V. 2011. Mapping State Proficiency Standards Onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005–2009. NCES 2011-458. *National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC: Government Printing Office.*

Bandeira de Mello, V., G. Bohrnstedt, C. Blankenship and D. Sherman. 2015. Mapping State Proficiency Standards Onto NAEP Scales: Results From the 2013 NAEP Reading and Mathematics Assessments. NCES 2015-046. *U.S. Department of Education, Washington, DC: National Center for Education Statistics*. Retrieved Mar. 11, 2016 from http://nces.ed.gov/pubsearch.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 119(1), 249–75.

Bhatt, Rachana and Cory Koedel. 2012. Large-Scale Evaluations of Curricular Effectiveness: The Case of Elementary Mathematics in Indiana. *Educational Evaluation and Policy Analysis* 34(4), 391–412.

Botero, Juan C., Simeon Djankov, Rafael la Porta, Florencio Lopez-de-Silanes and Andrei Shleifer. 2004. The Regulation of Labor. *Quarterly Journal of Economics* 119(4), 1339–82.

Boyd, Donald, Pamela Grossman, Hamilton Lankford, Susanna Loeb and James Wyckoff. 2008. Measuring Effect Sizes: The Effect of Measurement Error. CALDER Working Paper No. 19.

Branch, Gregory F., Eric A. Hanushek and Steven G. Rivkin. 2012. Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals. NBER Working Paper No. 17803.

Burgess, Simon, Deborah Wilson and Jack Worth. 2013. A Natural Experiment in School Accountability. The Impact of School Performance Information on Pupil Progress. *Journal of Public Economics* 106, 57-67.

Caballero, Ricardo J., Kevin N. Cowan, Eduardo M.R.A Engel and Alejandro Micco. 2013. Effective Labor Regulation and Microeconomic Flexibility. *Journal of Development Economics* 101, 92-104.

Cannata, Marissa. 2010. Understanding the Teacher Job Search Process: Espoused Preferences and Preferences in Use. *Teachers College Record* 112(12), 2889-2934.

Chan, Sewin and Ann H. Stevens. 2008. What You Don't Know Can't Help You: Pension Knowledge and Retirement Decision-Making. *Review of Economics and Statistics* 90(2), 253–66.

Clark, Robert L., Melinda Sandler Morrill and Steven G. Allen. 2012. The Role of Financial Literacy in Determining Retirement Plans. *Economic Inquiry* 50(4), 851-866.

Coggshall, Jane G. and Susan K. Sexton. 2008. Teachers on the Move: A Look at Teacher Interstate Mobility Policy and Practice. *National Association of State Directors of Teacher Education and Certification (NJ1).*

Costrell, Robert M. and Michael Podgursky. 2010. Distribution of Benefits in Teacher Retirement Systems and Their Implications for Mobility. *Education Finance and Policy* 5(4), 519–57.

Cullen, Julie Berry, Brian A. Jacob and Steven D. Levitt. 2006. The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica* 74(5), 1191-1230.

Curran, Bridget, Camille Abrahams and Theresa Clarke. 2001. *Solving Teacher Shortages Through License Reciprocity*. Denver, CO: State Higher Education Executive Officers.

DePasquale, Christina and Kevin Stange. 2016. Labor Supply Effects of Occupational Regulation: Evidence from the Nurse Licensure Compact. NBER Working Paper No. 22344.

Domina, Thurston, Andrew McEachin, Andrew Penner and Emily Penner. 2014. Aiming High and Falling Short: California's Eighth-Grade Algebra-for-All Effort. *Educational Evaluation and Policy Analysis* 37(3), 275–95.

Engel, Mimi, Brian A. Jacob and F. Chris Curran. 2014. New Evidence on Teacher Labor Supply. *American Educational Research Journal* 51(1), 36-72.

Farber, Henry S. 1999. Mobility and Stability: The Dynamics of Job Change in Labor Markets in Handbook of Labor Economics Vol. 3B, eds. O. Ashenfelter and D. Card: 2439-2484. Amsterdam: Elsevier.

Fitzpatrick, Maria D., David Grissmer and Sarah Hastedt. 2011. What a Difference a Day Makes: Estimating Daily Learning Gains During Kindergarten and First Grade Using a Natural Experiment. *Economics of Education Review* 30(2), 269-279.

Fitzpatrick, Maria D. and Michael F. Lovenheim. 2014. Early Retirement Incentives and Student Achievement. *American Economic Journal: Economic Policy* 6(3), 120–54.

Goldhaber, Dan, Cyrus Grout, Kristian Holden and Nate Brown. 2015. Crossing the Border? Exploring the Cross-State Mobility of the Teacher Workforce. *Educational Researcher* 44(8), 421-431.

Greenberg, David and John McCall. 1974. Teacher Mobility and Allocation. *Journal of Human Resources* 9(4), 480-502.

Haltiwanger, John, Stefano Scarpetta and Helena Schweiger. 2006. Assessing the Job Flows Across Countries: The Role of Industry, Size and Regulations. Unpublished manuscript.

Hanushek, Eric A. and Steven G. Rivkin. 2010. Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review: Papers and Proceedings* 100(2), 267–71.

Helpman, Elhanan and Oleg Itskhoki. 2010. Labour Market Rigidities, Trade and Unemployment. *Review of Economic Studies* 77(3), 1100–1137.

Ho, Andrew D. 2008. The Problem With 'Proficiency': Limitations of Statistics and Policy Under No Child Left Behind. *Educational Researcher* 37(6), 351–60.

Jackson, C. Kirabo. 2013. Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *Review of Economics and Statistics* 95(4), 1096–1116.

Kleiner, Morris M. 2015. Reforming Occupational Licensing Policies. *The Hamilton Project*, January.

Koedel, Cory, Jason A. Grissom, Shawn Ni and Michael Podgursky. 2012. Pension-Induced Rigidites in the Labor Market for School Leaders. CALDER Working Paper No. 71.

Koedel, Cory and Michael Podgursky. 2016. "Teacher Pensions" in Eric A. Hanushek, Stephen Machin and Ludger Woessman (Eds.) *Handbook of Economics of Education*, volume 5, 281-304.

Lafontaine, Francine and Jagadeesh Sivadasan. 2009. Do Labor Market Rigidities have Microeconomic Effects? Evidence from Within the Firm. *American Economic Journal: Applied Economics* 1(2), 88-127.

Lefgren, Lars and David Sims. 2012. Using Subject Test Scores Efficiently to Predict Teacher Value-Added. *Educational Evaluation and Policy Analysis* 34(1), 109–21.

Miller, Raegen T., Richard J. Murnane and John B. Willett. 2008. Do Teacher Absences Impact Student Achievement? Longitudinal Evidence from One Urban School District. *Educational Evaluation and Policy Analysis* 30(2), 181-200.

Mitra, Devashish and Priya Ranjan. 2010. Offshoring and Unemployment: The Role of Search Frictions Labor Mobility. *Journal of International Economics* 81(2), 219–29.

Parsons, Eric, Cory Koedel, Michael Podgursky, Mark Ehlert and P. Brett Xiang. 2015. Incorporating End-of-Course Exam Timing into Educational Performance Evaluations. *Journal of Research on Educational Effectiveness* 8(1), 130-47.

Podgursky, Michael, Mark Ehlert, Jim Lindsay and Yinmae Wan. 2016. *Interstate and intrastate educator mobility in three Midwestern states*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.

Reardon, Sean, Demetra Kalogrides, Andrew Ho, Ben Shear, Kenneth Shores, Erin Fahle. 2016a. *Stanford Education Data Archive*. http://purl.stanford.edu/db586ns4974.

Reardon, Sean F., Shear, Ben R., Castellano, Katherine E., & Ho, Andrew D. 2016b. Using Heteroskedastic Ordered Probit Models to Recover Moments of Continuous Test Score Distributions from Coarsened Data (CEPA Working Paper No.16-02). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp16-02.

Sass, Tim R. 2015. Licensure and Worker Quality: A Comparison of Alternative Routes to Teaching. *Journal of Law and Economics* 58(1), 1–35.

Taylor, Eric and John H. Tyler. 2012. The Effect of Evaluation of Teacher Performance. *American Economic Review* 102(7), 3628–51.

Figure 1: Illustrative Example of the Construction of the Boundary Intensity Measure for Hypothetical School A

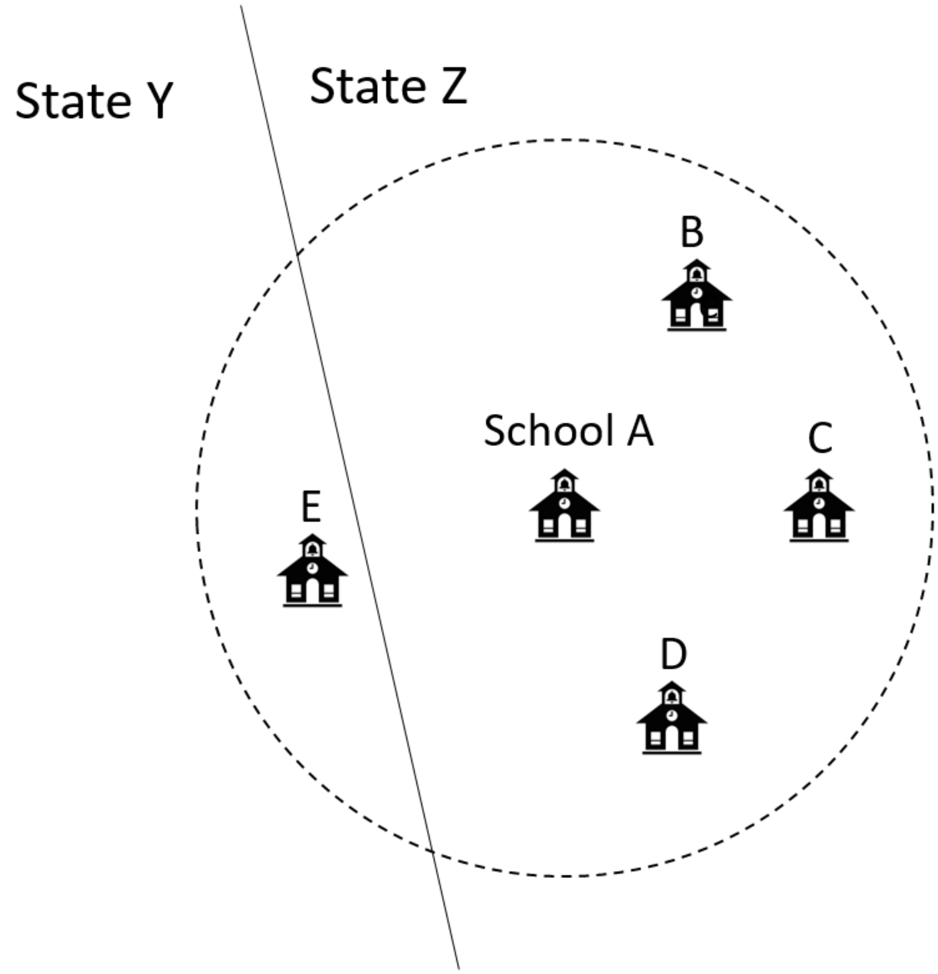
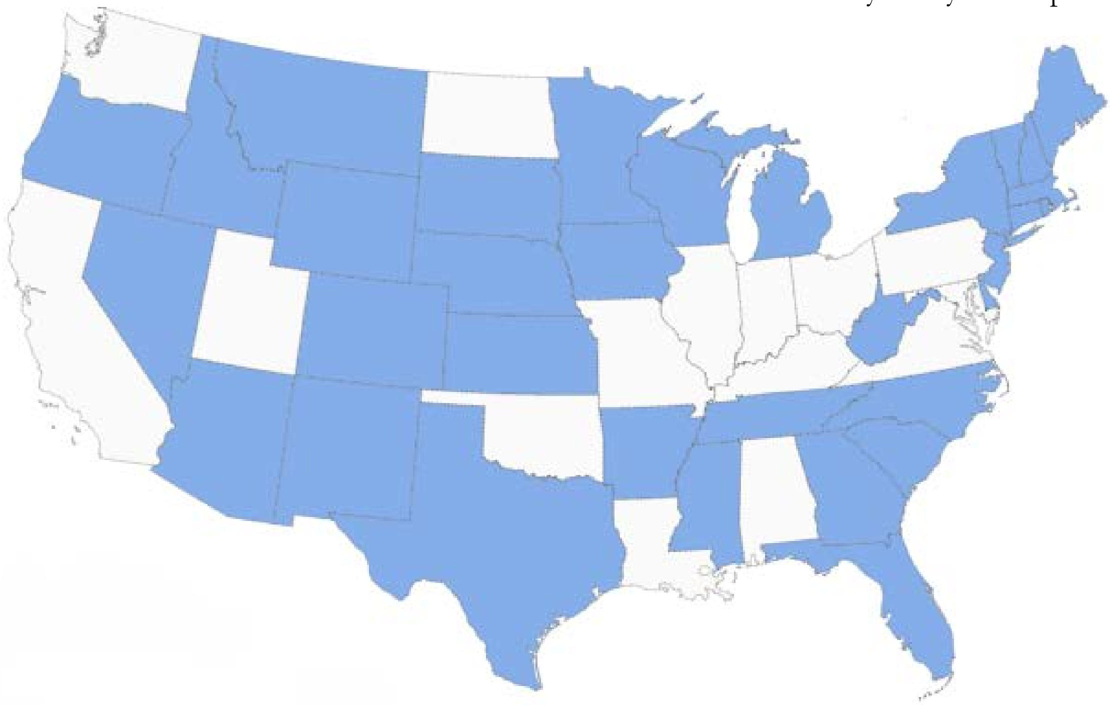


Figure 2: 33 States with Grade-8 Math Scaled Scores Included in the Primary Analytic Sample



Notes: The 33 states in the primary analytic sample for math are: Arkansas, Arizona, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Iowa, Kansas, Maine, Massachusetts, Michigan, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Vermont, West Virginia, Wisconsin, and Wyoming.

Table 1: Average Characteristics of Middle Schools in the CCD and Primary Analytic Sample

	All Schools in CCD		Primary Analytic Sample	
	Mean	St Dev	Mean	St Dev
Standardized Math Scaled Score	-	-	0.03	0.95
Standardized Reading Scaled Score	-	-	0.03	0.93
School Characteristics				
% Free Lunch Status	45.81	25.67	44.78	24.5
% Reduced Lunch Status	7.71	5.69	8.12	6.06
% White	58.46	33.65	61.02	32.16
% Black	14.71	24.15	13.89	22.25
% Hispanic	19.87	25.61	18.96	24.23
% Asian	3.13	6.78	2.37	4.77
% American Indian	1.28	6.88	1.59	8.11
% Pacific Islander	0.18	0.57	0.14	0.45
% Two or more races	2.36	2.76	2.04	2.21
Log of Total Enrollment	6.09	0.79	6.04	0.83
District Characteristics				
Log of Total District Enrollment	8.63	1.97	8.43	1.93
% English Language Learners	7.44	9.94	6.09	8
Log of Total Revenue per pupil	9.4	0.33	9.4	0.35
Log of Local Revenue per pupil	8.42	0.66	8.43	0.68
Zip Code Characteristics				
Log Median Household Income	10.82	0.37	10.8	0.37
% Low Education	45.82	15.21	45.53	14.62
Population Density	2170.02	4005.97	1595.06	3172.51
Urban-Centric Locale Categories				
Proportion of City Schools	22.73	41.91	19.4	39.54
Proportion of Suburb Schools	28.96	45.36	26.64	44.21
Proportion of Town Schools	12.92	33.54	13.58	34.26
Proportion of Rural Schools	35.39	47.82	40.38	49.07
Proportion by Labor Market Bifurcation				
Out-of-state Labor Market Percent $\geq$ 25	5.09	21.99	5.06	21.91
0 < Out-of-state Labor Market Percent < 25	7.12	25.72	6.18	24.08
N	18,396		11,686	

Notes: We use school records will full information to populate this table. The "Primary Analytic Sample" is the grade-8 math sample; the grade-8 reading sample includes California but excludes Nebraska due to testing issues as described in the text.

Table 2: Models of Selection into Boundary Regions

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample		Exclude "Moderately Affected" Schools		Out of State Labor Market Percent (Linear)	
VARIABLES	1(Out-of-State Labor Market Percent ≥ 25)					
% Free Lunch Status	0.0003	0.0002	0.0003	0.0002	0.0159	0.0068
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0125)	(0.0112)
% Reduced Lunch Status	-0.0004	0.0002	-0.0005	0.0004	-0.0181	0.0144
	(0.0005)	(0.0003)	(0.0006)	(0.0005)	(0.0301)	(0.0207)
% Asian	0.0016	0.0007	0.0017	0.0005	0.1074	0.0588
	(0.0022)	(0.0016)	(0.0024)	(0.0016)	(0.1398)	(0.1066)
% Hispanic	-0.0004	0.0001	-0.0004	0.0001	-0.0159	0.0196
	(0.0004)	(0.0003)	(0.0005)	(0.0003)	(0.0288)	(0.0195)
% Black	-0.0005	-0.0004*	-0.0005	-0.0003	-0.0186	-0.0123
	(0.0003)	(0.0002)	(0.0004)	(0.0002)	(0.0183)	(0.0124)
% American Indian	-0.0003	0.0001	-0.0004	0.0000	-0.0185	0.0086
	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0195)	(0.0178)
% Pacific Islander	0.0235*	0.0160	0.0243*	0.0147	1.0990*	0.6785
	(0.0127)	(0.0102)	(0.0132)	(0.0095)	(0.6021)	(0.4815)
% Two or more race	-0.0009	0.0006	-0.0010	0.0005	-0.0479	0.0475
	(0.0019)	(0.0011)	(0.0022)	(0.0012)	(0.1117)	(0.0638)
Log of Total Enrollment	0.0089	0.0061	0.0083	0.0051	0.3818	0.1045
	(0.0054)	(0.0040)	(0.0060)	(0.0043)	(0.3249)	(0.2276)
Log of Total District Enrollment	-0.0115	-0.0134	-0.0114	-0.0119*	-0.5844	-0.6996
	(0.0084)	(0.0081)	(0.0080)	(0.0065)	(0.4460)	(0.4297)
% English Language Learners	-0.0025	-0.0014	-0.0029	-0.0017	-0.1592	-0.0954
	(0.0016)	(0.0010)	(0.0019)	(0.0013)	(0.1005)	(0.0723)
Log of Total Revenue per pupil	0.0554*	-0.0009	0.0778*	0.0185	3.9634**	0.5601
	(0.0292)	(0.0235)	(0.0392)	(0.0160)	(1.7766)	(1.1088)
Log of Local Revenue per pupil	-0.0181***	-0.0089*	-0.0247**	-0.0184**	-1.2470***	-0.8956**
	(0.0064)	(0.0044)	(0.0098)	(0.0081)	(0.4290)	(0.3911)
Log Median Household Income	0.0577*	0.0222	0.0715*	0.0276	3.6176**	1.4831*
	(0.0293)	(0.0142)	(0.0369)	(0.0171)	(1.5199)	(0.8721)
% Low Education	0.0016**	0.0005	0.0019**	0.0006	0.0963**	0.0227
	(0.0007)	(0.0003)	(0.0009)	(0.0004)	(0.0364)	(0.0165)
Population Density/1000	0.0096	0.0064	0.0129	0.0095	0.7188*	0.5294
	(0.0060)	(0.0052)	(0.0086)	(0.0073)	(0.4027)	(0.3677)
1(Suburb)	0.0004	-0.0100	0.0065	-0.0078	0.3332	-0.3904
	(0.0092)	(0.0123)	(0.0102)	(0.0116)	(0.5481)	(0.6997)
1(Town)	-0.0214	-0.0287*	-0.0179	-0.0265*	-0.6536	-1.1340
	(0.0142)	(0.0145)	(0.0156)	(0.0142)	(0.8375)	(0.7821)
1(Rural)	-0.0357*	-0.0342*	-0.0350*	-0.0336*	-1.5418	-1.5063
	(0.0184)	(0.0184)	(0.0198)	(0.0180)	(1.0001)	(0.9668)
Constant	-0.9563**	-0.1058	-1.2686**	-0.2774	-64.1332***	-10.3452
	(0.3797)	(0.2574)	(0.5465)	(0.2620)	(19.6322)	(17.4594)
State Fixed Effects		X		X		X
R-squared	0.0540	0.1164	0.0704	0.1576	0.0777	0.1458
Observations (schools)	11,686	11,686	10,964	10,964	11,686	11,686
Joint P-Value		0.263		0.264		0.691

Notes: This table shows variants of the selection equation in the main text where we adjust the dependent variable and/or the sample. The first four columns predict an indicator for being an intensely-affected boundary school using available covariates. Columns 1 and 2 assign values of 0 for the indicator for all other schools; columns 3 and 4 drop moderately affected schools (with out-of-state FTE shares above 0 but less than 0.25). Columns 5 and 6 are from a model that includes all schools where the dependent variable is the out-of-state FTE share (linear). Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimated Boundary Effects on Grade 8 Scaled Scores

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.0942**	-0.0537*
	(0.0377)	(0.0283)
0 < Out-of-State Labor Market Percent < 25	-0.0095	0.0082
	(0.0330)	(0.0189)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4773	0.5896
Observations (schools)	11,686	13,286

Notes: The difference in the number of schools between the math and reading models is primarily due to a one-state difference in the samples (see footnote 7); there are also a small number of schools that report one score or the other but they are so few as to be inconsequential. Coefficients for non-boundary coefficients are reported in Appendix Table A4. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Estimated Boundary Effects on the Grade 8 Proficiency Rate

	(1)	(2)	(3)	(4)
	Scaled Score Sample		Extended Sample (43 states)	
	Grade 8 Proficiency Rate			
VARIABLES	Math	Reading	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.0816**	-0.0329	-0.0772***	-0.0319
	(0.0331)	(0.0295)	(0.0285)	(0.0376)
0 < Out-of-State Labor Market Percent < 25	-0.0082	0.0412	0.0160	0.0761**
	(0.0384)	(0.0298)	(0.0399)	(0.0297)
Covariates	X	X	X	X
State Fixed Effects	X	X	X	X
R-squared	0.4354	0.5242	0.4288	0.5070
Observations (schools)	11,512	13,180	16,269	18,001

Notes: Columns 1 and 2 use the same states with scaled score data from Table 3 (33 states). Columns 3 and 4 use all states where proficiency rate data are available (43 states). The small sample-size differences between columns 1 and 2 here, and Table 3, are because scale scores and proficiency rates are not both available for all schools. Standard errors are clustered at the state level.

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

Table 5: Robustness Results, Including 20-mile Radius Measures

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
10-mile radius Out-of-State Labor Market Percent ≥ 25	-0.0742	-0.0632*
	(0.0448)	(0.0382)
0 < 10-mile radius Out-of-State Labor Market Percent < 25	-0.0055	-0.0029
	(0.0373)	(0.0266)
20-mile radius Out-of-State Labor Market Percent ≥ 25	-0.0219	0.0143
	(0.0360)	(0.0336)
0 < 20-mile radius Out-of-State Labor Market Percent < 25	0.0315	0.0188
	(0.0324)	(0.0242)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4775	0.5896
Observations (schools)	11,686	13,286

Notes: Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Robustness Results, Estimated Using Counts of In-State and Out-of-State Local FTE

	(1)	(2)	(3)	(4)
	Grade 8 Scaled Score			
VARIABLES	Math	Reading	Math	Reading
Total FTE10/1000	0.0128**	0.0214***	0.0140	0.0251***
	(0.0062)	(0.0066)	(0.0084)	(0.0089)
$(\text{Total FTE10/1000})^2$	-0.0002	-0.0003***	0.0001	-0.0003
	(0.0001)	(0.0001)	(0.0003)	(0.0004)
Out-of-State FTE10/1000			-0.0385***	-0.0364***
			(0.0081)	(0.0092)
$(\text{Out-of-State FTE10/1000})^2$			0.0006	0.0009*
			(0.0006)	(0.0006)
Covariates	X	X	X	X
State Fixed Effects	X	X	X	X
R-squared	0.4775	0.5911	0.4781	0.5917
Observations (schools)	11,686	13,286	11,686	13,286

Notes: Total FTE10/1000 and Out-of-State FTE10/1000 are the total number of local FTE and out-of-state local FTE (in the circle with radius 10 miles), respectively, divided by 1000. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Robustness Results, Heterogeneous Effects Estimated by Population Density

VARIABLES	(1) Primary Model	(2) Population-Density Interactions
Math - Grade 8 Scaled Score		
Out-of-State Labor Market Percent ≥ 25	-0.0942**	-0.0982**
	(0.0377)	(0.0377)
0 < Out-of-State Labor Market Percent < 25	-0.0095	-0.0158
	(0.0330)	(0.0305)
Out-of-State Labor Market Percent ≥ 25 *Below Median		0.0081
		(0.0751)
0 < Out-of-State Labor Market Percent < 25 *Below Median		0.0233
		(0.0557)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4773	0.4773
Observations (schools)	11,686	11,686
Reading - Grade 8 Scaled Score		
Out-of-State Labor Market Percent ≥ 25	-0.0537*	-0.0731**
	(0.0283)	(0.0323)
0 < Out-of-State Labor Market Percent < 25	0.0082	0.0038
	(0.0189)	(0.0215)
Out-of-State Labor Market Percent ≥ 25 *Below Median		0.0438
		(0.0681)
0 < Out-of-State Labor Market Percent < 25 *Below Median		0.0108
		(0.0416)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.5896	0.5896
Observations (schools)	13,286	13,286

Notes: Standard errors are clustered at the state level. We obtain qualitatively similar results if we divide the sample by urbanicity designation (urban and suburban versus rural and town) instead of population density.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Robustness Results, Estimated Using Grade-7, Grade-5 and Grade-3 Scaled Scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Grade 7		Grade 5		Grade 3	
VARIABLES	Math	Reading	Math	Reading	Math	Reading
Out-of-State Percent ≥ 25	-0.0696*	-0.0417	-0.0459	-0.0396	-0.0423	0.0056
	(0.0378)	(0.0341)	(0.0400)	(0.0323)	(0.0421)	(0.0390)
0 < Out-of-State Percent < 25	0.0073	0.0369	-0.0241	0.0065	0.0112	-0.0132
	(0.0391)	(0.0250)	(0.0550)	(0.0354)	(0.0392)	(0.0356)
Covariates	X	X	X	X	X	X
State Fixed Effects	X	X	X	X	X	X
R-squared	0.4945	0.5890	0.5000	0.6310	0.5041	0.6021
Observations (schools)	13,878	13,631	28,965	28,499	29,000	28,859

Notes: The grade-7 math sample is substantially larger than the grade-8 math sample (from Table 3) because we include California. The grade-7 reading sample varies slightly from the grade-8 sample because of small differences in which schools report scores for which grades. The elementary school sample sizes are much larger because there are many more elementary schools than middle schools in the data. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Robustness Results, Estimated Using District-level Performance Metrics from the Stanford Education Data Archive

	(1)	(2)	(3)	(4)
	Primary Scaled-Score Sample		Extended Sample (all lower-48 states)	
VARIABLES	District-level Grade 8 Scaled Score			
	Math	Reading	Math	Reading
Share of "Intensely Affected" Boundary Schools	-0.0401**	-0.0192	-0.0430**	-0.0260**
	(0.0175)	(0.0128)	(0.0166)	(0.0104)
Share of "Moderately Affected" Boundary Schools	-0.0062	0.0100	0.0005	0.0040
	(0.0206)	(0.0167)	(0.0167)	(0.0125)
Covariates	X	X	X	X
State Fixed Effects	X	X	X	X
R-squared	0.4769	0.5994	0.4485	0.5677
Observations (districts)	6,087	6,710	9,549	10,346

Notes: Columns 1 and 2 use states with scaled score data from Table 3 (33 states). Column 3 uses all lower-48 states except California and column 4 uses all lower-48 states. SEDA does not provide grade-8 math data from California for the same reason that we do not include these data (see text). Intensely affected schools are defined as those with 25 percent or more of the local-area labor market on the other side of a state line. Moderately affected schools are defined as those with more than zero but less than 25 percent of the local-area labor market on the other side of a state line. Standard errors are clustered at the state level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

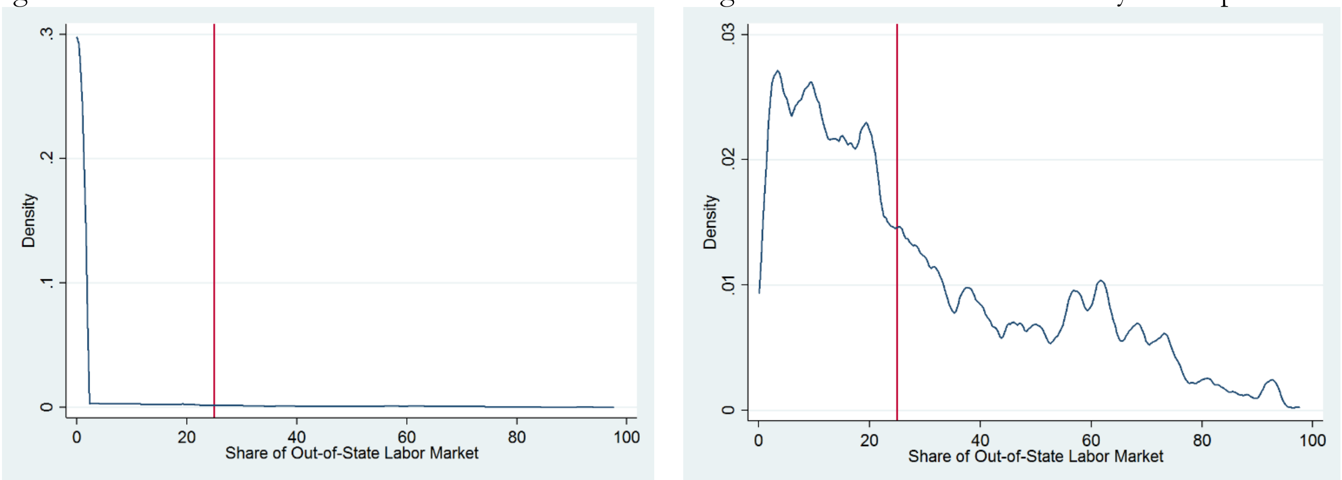
Table 10: Models that Include Out-of-District Labor Market Measures

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.0852**	-0.0471
	(0.0369)	(0.0281)
0 < Out-of-State Labor Market Percent < 25	-0.0047	0.0130
	(0.0324)	(0.0185)
Out-of-District Labor Market Percent ≥ 25	-0.0595***	-0.0541***
	(0.0184)	(0.0175)
0 < Out-of-District Labor Market Percent < 25	-0.0420	-0.0574*
	(0.0331)	(0.0327)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4777	0.5899
Observations (schools)	11,686	13,286

Notes: The "out-of-district labor market percent" variable uses a 10-mile radius to define a local area labor market as with the out-of-state measure. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

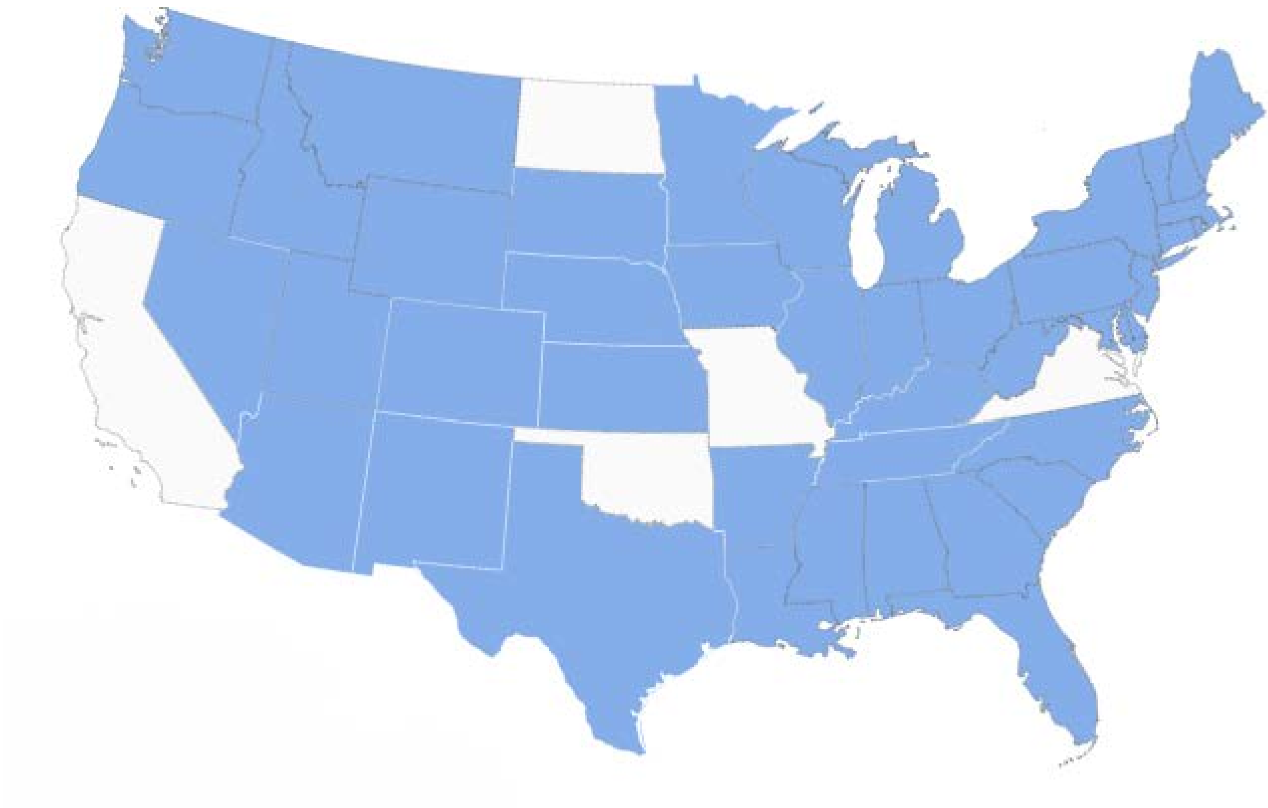
Appendix A
Supplementary Figures and Tables

Figure A1: Distributions of the Out-of-State FTE Percentage for Schools in the Math Analytic Sample



Notes: The left panel shows the full distribution of out-of-state FTE percentages, which is dominated by the 88 percent of schools that have zero out-of-state FTE. The right panel shows the distribution of out-of-state FTE percentages conditional on a non-zero value. The vertical line shows the 25-percent FTE cutoff that we use in our preferred measure of boundary intensity. As noted in the text, even conditional on a non-zero out of state share, most schools have a larger percentage of in-state FTE. The reason is that each school's circle is centered on itself, which in most cases will lead to the in-state area being larger than the out-of-state area. Kernel density estimates use the Epanechnikov kernel function with a kernel width of two.

Figure A2: 43 States with Grade-8 Math Proficiency Rate Data Included in the Extended Sample



Notes: List of 43 states in the extended sample for math are: Arkansas, Alabama, Arizona, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, West Virginia, Washington, Wisconsin, and Wyoming.

Table A1: Sample Construction Details

	2012-2013 Grade-8 Test Results							
	Scaled Score				Proficiency Rate			
	Math		Reading		Math		Reading	
	Schools	% of Universe	Schools	% of Universe	Schools	% of Universe	Schools	% of Universe
Universe	12,992		14,705		18,200		20,103	
Missing Data								
CCD School	-262	2.02%	-282	1.92%	-330	1.81%	-370	1.84%
CCD District Enroll	-458	3.53%	-459	3.12%	-857	4.71%	-869	4.32%
CCD District Fin.	-76	0.58%	-96	0.65%	-69	0.38%	-96	0.48%
ACS ZIP-Code	-89	0.69%	-93	0.63%	-115	0.63%	-121	0.60%
Geocoded Data	-421	3.24%	-489	3.33%	-560	3.08%	-646	3.21%
Final Sample	11,686	89.95%	13,286	90.35%	16,269	89.39%	18,001	89.54%
Number of States	33		33		43		43	

Notes: No test results from Missouri, North Dakota, or Oklahoma are included in the table. Reading test results are not available for Nebraska and we do not use California math scores for reasons discussed in the text. Scaled scores are not available for Alabama, Illinois, Indiana, Kentucky, Louisiana, Maryland, Ohio, Pennsylvania, Utah, and Washington.

Table A2: Average Characteristics of Middle Schools in the CCD and Proficiency Rate Sample

	All Schools in CCD		Proficiency Rate Sample	
	Mean	St Dev	Mean	St Dev
Standardized Math Proficiency Rate	-	-	0.04	0.95
Standardized Reading Proficiency Rate	-	-	0.05	0.92
School Characteristics				
% Free Lunch Status	45.81	25.67	44.93	25.36
% Reduced Lunch Status	7.71	5.69	7.68	5.72
% White	58.46	33.65	61.7	32.78
% Black	14.71	24.15	15.57	24.96
% Hispanic	19.87	25.61	16.62	23.06
% Asian	3.13	6.78	2.42	4.94
% American Indian	1.28	6.88	1.25	6.88
% Pacific Islander	0.18	0.57	0.14	0.48
% Two or more races	2.36	2.76	2.3	2.5
Log of Total Enrollment	6.09	0.79	6.08	0.76
District Characteristics				
Log of Total District Enrollment	8.63	1.97	8.57	1.95
% English Language Learners	7.44	9.94	5.77	7.68
Log of Total Revenue per pupil	9.4	0.33	9.42	0.32
Log of Local Revenue per pupil	8.42	0.66	8.47	0.64
Zip Code Characteristics				
Log Median Household Income	10.82	0.37	10.8	0.37
% Low Education	45.82	15.21	46.16	14.86
Population Density	2170.02	4005.97	1967.44	3833.43
Urban-Centric Locale Categories				
Proportion of City Schools	22.73	41.91	21.4	41.01
Proportion of Suburb Schools	28.96	45.36	27.88	44.84
Proportion of Town Schools	12.92	33.54	13.55	34.23
Proportion of Rural Schools	35.39	47.82	37.17	48.33
Proportion by Labor Market Bifurcation				
Out-of-state Labor Market Percent ≥ 25	5.09	21.99	5.66	23.11
0 < Out-of-state Labor Market Percent < 25	7.12	25.72	7.95	27.05
Observations (schools)	18,396		16,269	

Notes: We use school records will full information to populate this table.

Table A3: Additional Models of Selection into Boundary Regions

VARIABLES	(1)	(2)	(3)	(4)
	1(Out-of-State Labor Market Percent > 0)		Exclude "Moderately Affected" Schools and Use More-Interior Control Group	
% Free Lunch Status	0.0005	-0.0000	0.0003	-0.0001
	(0.0004)	(0.0002)	(0.0004)	(0.0002)
% Reduced Lunch Status	-0.0002	0.0006	-0.0009	-0.0000
	(0.0012)	(0.0008)	(0.0009)	(0.0004)
% Asian	0.0026	0.0014	0.0027	0.0006
	(0.0024)	(0.0015)	(0.0034)	(0.0017)
% Hispanic	-0.0006	0.0008	-0.0004	0.0001
	(0.0007)	(0.0005)	(0.0007)	(0.0003)
% Black	-0.0004	0.0002	-0.0005	-0.0002
	(0.0007)	(0.0005)	(0.0006)	(0.0003)
% American Indian	-0.0004	0.0007	-0.0005	0.0002
	(0.0006)	(0.0005)	(0.0005)	(0.0004)
% Pacific Islander	0.0241	0.0135	0.0256*	0.0116
	(0.0145)	(0.0100)	(0.0140)	(0.0088)
% Two or more race	-0.0005	0.0015	-0.0014	0.0007
	(0.0039)	(0.0025)	(0.0030)	(0.0014)
Log of Total Enrollment	0.0074	-0.0048	0.0121	0.0056
	(0.0113)	(0.0074)	(0.0075)	(0.0049)
Log of Total District Enrollment	-0.0060	-0.0049	-0.0135	-0.0108*
	(0.0098)	(0.0074)	(0.0096)	(0.0054)
% English Language Learners	-0.0042*	-0.0027*	-0.0035	-0.0014
	(0.0021)	(0.0016)	(0.0022)	(0.0010)
Log of Total Revenue per pupil	0.1717***	0.0528	0.1222*	0.0227
	(0.0600)	(0.0340)	(0.0642)	(0.0194)
Log of Local Revenue per pupil	-0.0431**	-0.0314	-0.0362**	-0.0193*
	(0.0186)	(0.0213)	(0.0168)	(0.0110)
Log Median Household Income	0.1342**	0.0645	0.1032*	0.0345
	(0.0534)	(0.0438)	(0.0560)	(0.0240)
% Low Education	0.0028**	0.0008	0.0026*	0.0007
	(0.0012)	(0.0006)	(0.0014)	(0.0004)
Population Density/1000	0.0171*	0.0104	0.0148	0.0087
	(0.0089)	(0.0085)	(0.0107)	(0.0080)
1(Suburb)	-0.0007	-0.0201	0.0178	0.0042
	(0.0214)	(0.0200)	(0.0181)	(0.0094)
1(Town)	-0.0413	-0.0467*	-0.0174	-0.0195
	(0.0326)	(0.0268)	(0.0198)	(0.0155)
1(Rural)	-0.0772**	-0.0643**	-0.0393*	-0.0264
	(0.0347)	(0.0266)	(0.0227)	(0.0157)
Constant	-2.6826***	-0.8847	-1.9497**	-0.3932
	(0.6378)	(0.6855)	(0.9173)	(0.3686)
State Fixed Effects		X		X
R-squared	0.1052	0.1997	0.0993	0.2646
Observations (schools)	11,686	11,686	9,434	9,434
Joint p-value		0.444		0.702

Notes: This table provides additional selectivity tests following on Table 2 in the main text. Columns 1 and 2 code the dependent variable to one for schools with any non-zero out-of-state FTE share and zero otherwise; columns 3 and 4 code the dependent variable to one for intensely affected schools and zero otherwise, but further differentiate intensely-affected schools from control schools by dropping all control schools with any out-of-state FTE share within a *20-mile* radius. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Estimated Coefficients for Non-boundary Controls from Main Models in Table 3

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
% Free Lunch Status	-0.0178***	-0.0189***
	(0.0024)	(0.0024)
% Reduced Lunch Status	-0.0070***	-0.0049
	(0.0019)	(0.0034)
% Asian	0.0202***	0.0118***
	(0.0037)	(0.0012)
% Hispanic	-0.0048***	-0.0032***
	(0.0014)	(0.0010)
% Black	-0.0113***	-0.0106***
	(0.0014)	(0.0014)
% American Indian	-0.0129***	-0.0101***
	(0.0023)	(0.0017)
% Pacific Islander	-0.0039	-0.0359**
	(0.0124)	(0.0140)
% Two or more race	-0.0058	-0.0035
	(0.0049)	(0.0028)
Log of School Enrollment	0.2441***	0.2692***
	(0.0439)	(0.0377)
Log of District Enrollment	-0.0703***	-0.0870***
	(0.0173)	(0.0101)
% English Language Learners	0.0088**	0.0033**
	(0.0041)	(0.0015)
Log of Total District Revenue Per Pupil	0.0762	-0.1427
	(0.0890)	(0.1596)
Log of Local District Revenue Per Pupil	-0.0012	0.0179
	(0.0357)	(0.0351)
Log Median Household Income	-0.0822	-0.1367***
	(0.0508)	(0.0482)
% Low Education	-0.0068***	-0.0114***
	(0.0015)	(0.0012)
Population Density/1000	0.0123**	0.0112**
	(0.0046)	(0.0050)
1(Suburb)	0.0368	0.0352
	(0.0309)	(0.0209)
1(Town)	0.0044	-0.0218
	(0.0398)	(0.0293)
1(Rural)	0.0520	0.0608*
	(0.0332)	(0.0324)
Constant	0.6889	3.2772**
	(0.7559)	(1.2419)
State Fixed Effects	X	X
R-squared	0.4773	0.5896
Observations (schools)	11,686	13,286

Notes: Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table A5: Replication of Results in Table 3 Using Sparse Control Sets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Grade 8 Scaled Score Math				Grade 8 Scaled Score Reading			
Out-of-State Percent ≥ 25	-0.1060	-0.0897**	-0.0862**	-0.0942**	-0.0738	-0.0464*	-0.0455	-0.0537*
	(0.0807)	(0.0356)	(0.0350)	(0.0377)	(0.0904)	(0.0274)	(0.0276)	(0.0283)
0 <Out-of-State Percent< 25	-0.1303*	-0.0182	-0.0032	-0.0095	-0.1120	0.0028	0.0191	0.0082
	(0.0719)	(0.0377)	(0.0358)	(0.0330)	(0.0773)	(0.0200)	(0.0196)	(0.0189)
State Fixed Effects	X	X	X	X	X	X	X	X
School Controls		X	X	X		X	X	X
District Controls			X	X			X	X
Local Community Controls				X				X
R-squared	0.0047	0.4672	0.4727	0.4773	0.0076	0.5724	0.5775	0.5896
Observations (schools)	11,686	11,686	11,686	11,686	13,286	13,286	13,286	13,286

Notes: Standard errors are clustered at the state level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A6: Estimated Boundary Effects on Grade 8 Scaled Scores, Dropping Schools with No Other Schools within 10 miles

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.0946**	-0.0519*
	(0.0379)	(0.0285)
0 < Out-of-State Labor Market Percent < 25	-0.0098	0.0090
	(0.0330)	(0.0189)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4833	0.5938
Observations (schools)	11,457	13,052

Notes: The sample sizes in this table are smaller than in Table 3 because we drop schools without any other schools in the 10-mile radius. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix B
Sensitivity Analysis

The purpose of this appendix is to examine the robustness of our findings to a variety of ways of measuring boundary exposure.

First, in Table B1 we examine the sensitivity of our findings to alternative ways of constructing our measure of boundary exposure, all following the general structure of Equation (2) in the main text. The estimates in Table B1 are comparable to our primary findings in Table 3. The first two columns restrict the exposure measure to only include schools that have overlapping grades with the reference school. The middle two columns consider a similar, alternative restriction where we construct the analogs to Equation (2) restricted to schools where the share of free/reduced-price eligible students is within 20 percent of the share at the reference school. The last two columns return to the full sample of local-area schools, but replace the FTE-based measures in Equation (2) with analogous measures based on school enrollment. In all cases the substance of our results is maintained, although in the middle columns the results are weaker.

Next, in Table B2 we show results from models where we further differentiate boundary schools from control schools. To do this, we drop from the control group all schools where there is at least one other school *within 20 miles* that is on the other side of a state line. Thus, the control group in Table B2 is even less likely to be influenced by a state boundary.[20] The estimates in Table B2 are similar to what we show in Table 3, which is consistent with the locality of the boundary effect documented throughout our study (i.e., schools moderately far from a state boundary are no more affected than schools that are much further away).

Tables B3 and B4 show results from models that include linear and quadratic measures of boundary exposure. In Table B3, the linear and quadratic terms are for the out-of-state FTE share,

[20] This restricted control group is also examined in the analysis of selection into boundary regions in Appendix Table A3.

which is the same variable that we use to divide schools into "bins" for our main analysis. In Table B4, we use linear and quadratic terms that capture of the simple distance to the nearest state boundary as the crow flies. In both sets of estimates, our findings are substantively maintained. The models based on the distance to a state line show a pattern of statistical significance between math and reading that is reversed relative to other estimates throughout our paper, but we do not put much weight on this result given its uniqueness across the variety of robustness and sensitivity tests we consider.

Finally, in Table B5 we return to the structure of our primary measures of boundary exposure (i.e., the "binned" intensity variables), but adjust Equation (2) to include each school's own FTE in the denominator. Although our preferred construct omits own-school FTE from Equation (2), it is not unreasonable to include it. Table B5 shows that our findings are not substantively sensitive to including the own-school FTE share directly in the measure.[21]

[21] Note that our standard errors increase in Table B5, in part because a smaller fraction of schools meets the pre-defined cutoffs for the boundary-exposure bins. We can better connect the estimates in Table B5 to our estimates in Table 3 by adjusting the thresholds for the bins to maintain a constant share of the sample coded to each level of intensity. We omit results from this extended sensitivity analysis for brevity but they are substantively similar to what we show in Table B5.

Table B1: Robustness Results, Estimated Using Different Measures of the Intensity of Boundary Exposure

	(1)	(2)	(3)	(4)	(5)	(6)
	FTE in Schools with Overlapping Grade		FTE in Schools with Similar FRL Percent		School Enrollment	
	Grade 8 Scaled Scores					
VARIABLES	Math	Reading	Math	Reading	Math	Reading
Out-of-State Percent ≥ 25	-0.0932**	-0.0596*	-0.0522	-0.0016	-0.0864**	-0.0388
	(0.0420)	(0.0327)	(0.0344)	(0.0312)	(0.0349)	(0.0267)
0 < Out-of-State Percent < 25	-0.0155	-0.0072	0.0040	0.0040	-0.0172	-0.0068
	(0.0429)	(0.0235)	(0.0346)	(0.0202)	(0.0330)	(0.0189)
Covariates	X	X	X	X	X	X
State Fixed Effects	X	X	X	X	X	X
R-squared	0.4773	0.5896	0.4770	0.5894	0.4773	0.5895
Observations (schools)	11,686	13,286	11,686	13,286	11,686	13,286

Notes: The "FTE in Schools with Overlapping Grade" measure uses FTE only in local schools that share an overlapping grade with the reference school. The "FTE in Schools with Similar FRL Percent" measure uses FTE only in local schools that have a similar share of free/reduced-price lunch eligible students to the reference school (i.e., +/- 20 percent). The "School Enrollment" measure uses school enrollment in local schools instead of FTE to capture boundary exposure. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table B2: Robustness Results, Estimated Using a Restricted Control Group of Non-Boundary Schools Without Any Out-of-State FTE within a 20-mile Radius

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.1063**	-0.0581*
	(0.0427)	(0.0291)
0 < Out-of-State Labor Market Percent < 25	-0.0042	0.0148
	(0.0376)	(0.0207)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4795	0.5903
Observations (schools)	10,156	11,791

Notes: The sample sizes in this table are smaller than in Table 3 because the restricted control group is limited to schools that have zero out-of-state FTE within the 20-mile radius. Standard errors are clustered at the state level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B3: Robustness Results, Estimated with Linear and Quadratic Controls for the Out-of-State FTE Share

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent	-0.0031*	-0.0017
	(0.0018)	(0.0017)
$(\text{Out-of-State Labor Market Percent})^2$	0.0000	0.0000
	(0.0000)	(0.0000)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4772	0.5895
Observations (schools)	11,686	13,286

Notes: Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table B4: Robustness Results, Estimated with Linear and Quadratic Controls for the Distance to the Nearest State Boundary (in miles), Irrespective of Local Area FTE

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Distance to the Closest State Boundary/100	-0.0373	-0.1114*
	(0.0581)	(0.0566)
$(\text{Distance to the Closest State Boundary}/100)^2$	0.0046	0.0418
	(0.0124)	(0.0285)
State Fixed Effects	X	X
R-squared	0.4771	0.5903
Observations (schools)	11,686	13,286

Notes: Distance to the closest state boundary is measured in miles as the crow flies. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1

Table B5: Robustness Results, Estimated Using a Measure of Boundary Exposure with Own School FTE in the Denominator

	(1)	(2)
	Grade 8 Scaled Score	
VARIABLES	Math	Reading
Out-of-State Labor Market Percent ≥ 25	-0.0886**	-0.0476
	(0.0430)	(0.0327)
0 < Out-of-State Labor Market Percent < 25	-0.0252	-0.0069
	(0.0331)	(0.0191)
Covariates	X	X
State Fixed Effects	X	X
R-squared	0.4773	0.5895
Observations (schools)	11,686	13,286

Notes: The local-area labor market percent is calculated with the reference school's FTE in the denominator. Standard errors are clustered at the state level.
*** p<0.01, ** p<0.05, * p<0.1