Accepted Manuscript


# Labor Market Frictions and Production Efficiency in Public Schools

Dongwoo Kim , Cory Koedel , Shawn Ni , Michael Podgursky

PII: S0272-7757(16)30694-X
DOI: 10.1016/j.econedurev.2017.07.009
Reference: ECOEDU 1727

To appear in: *Economics of Education Review*

Received date: 16 December 2016
Revised date: 14 July 2017
Accepted date: 17 July 2017

Please cite this article as: Dongwoo Kim , Cory Koedel , Shawn Ni , Michael Podgursky , Labor Market Frictions and Production Efficiency in Public Schools, *Economics of Education Review* (2017), doi: 10.1016/j.econedurev.2017.07.009

## Highlights

- 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 using geocoded data on school locations and 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 measured dimensions.
- Achievement is lower in mathematics, and to a lesser extent in reading, at schools that are more exposed to state boundaries.

# Labor Market Frictions and Production Efficiency in Public Schools

Dongwoo Kim
Cory Koedel
Shawn Ni
Michael Podgursky

Kim, Ni and Podgursky are in the Department of Economics, and Koedel is in the Department of Economics and Truman School of Public Affairs, at the University of Missouri, Columbia.

July 2017

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 geocoded data on school locations and state boundaries. We find that achievement is lower in mathematics, and to a lesser extent in reading, at schools that are more 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 mobility frictions in educator labor markets near state boundaries 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 placements on salary schedules 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 large literature examining restricted labor mobility in other sectors points 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 information on schools in the United States merged with achievement data to empirically test whether exposure to state boundaries reduces schooling output. We find evidence

[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.

of a highly-localized, robust negative effect 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 line 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 a negative boundary effect on reading test scores, but it is smaller than in math and 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. Although our models suggest that there may be costly frictions associated with district lines within states, district frictions cannot explain the state-boundary effects.

## 2. Background

In this section we briefly discuss state-specific licensing and pension policies that impose additional costs on educator mobility across state lines. These policies motivate our examination of state boundary effects on student achievement; in the conclusion we also discuss other factors associated with state boundaries that may introduce labor frictions and contribute to our findings.

### *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. In addition to variation in initial requirements across states, there is also variation in what types of licenses cover what work. Cogshall and Sexton (2008) give the example that in some states special education teachers are licensed to teach children with a specific disability, while in others the license applies to students with any disability. Another example is that some states have dedicated licenses for middle school teachers but in others, middle school teachers are covered under a broader licensing category for teachers in grades 6-12.

How teachers progress through the levels of licensure within states – e.g., from a "level 1" to "level 2" license (in Vermont, for example, these are labeled the "initial" and "professional" levels, respectively) – also varies across states. First, in terms of structure, states differ in the number of licensing levels a teacher can obtain. Coggshall and Sexton (2008) document that twelve states have just one licensing tier, nineteen states and the District of Columbia have two tiers, and nineteen states have three tiers. Among states with multiple licensing tiers, there is cross-state variability in the labeling of different tiers, as well as in the substantive requirements to "move up." The requirements typically include combinations of experience, professional development and coursework, performance-based assessments and minimum scores on licensure tests.

The Interstate Agreement (IA), created by the National Association of State Directors of Teacher Education and Certification (NASDTEC), reflects the policy concern that state-specific licensing requirements restrict educator labor flows. The IA includes individual agreements between most US states outlining the processes for obtaining a license for transfers (the IA includes individual agreements among 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), which reflects cross-state heterogeneity in licensing rules. Our review of the IA suggests that for many states the required steps are substantial (e.g., taking specific tests, completing new coursework, etc.). Moreover, even in cases where reciprocity is complete or nearly so, the general complexity of state 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]

DePasquale and Stange (2016) find that a reduction in licensing barriers for nurses brought on by the Nurse Licensure Compact (NLC) did not increase cross-state labor mobility. One interpretation of their findings is that licensing barriers are unimportant, at least in the market for nurses, but there are several caveats to this interpretation. First, 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. DePasquale and Stange (2016) also cannot rule out fairly large mobility effects of the NLC relative to the baseline mobility rate in some specifications. Finally, their analysis does not isolate mobility near state boundaries beyond looking at boundary-touching counties, which can cover large geographic areas. Recent evidence on workers' strong preferences for short commutes suggests that mobility effects will be most pronounced very close to boundaries (Manning and Petrongolo, forthcoming).

### *2.2 Teacher Pensions*

State-specific pension coverage is another source of cross-state mobility costs for public educators. Most teachers are enrolled in 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.

[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.

There are two 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 value of the pension. 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 effects of state-boundary exposure on student achievement using linear regression models of the following form:

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

In Equation (1), $Y_{ij}$ is average achievement on the state standardized test for school $i$ in state $j$, normalized by state, grade, and subject.[3] 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 demographic and socioeconomic 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. $\gamma_j$ is a state fixed effect. $\varepsilon_{ij}$ is the error term, which we cluster at the state level.

The model in Equation (1) will yield unbiased estimates of boundary effects ($\boldsymbol{\delta_2}$) 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, below we show that there is no evidence of selection along any of the observed dimensions measured by the rich NCES and 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).

Our primary models use grade-8 test scores as outcomes. 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,

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

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, and if our findings are not driven by unobserved selection of schools into boundary regions, then we should estimate smaller effects in earlier grades. This 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 and the District of Columbia. We focus exclusively on traditional public schools.[4] We measure how much a school's local-area labor market is exposed to a state boundary by first drawing circles around the school with 10- and 20-mile radii. We then identify the total number of full-time equivalent (FTE) teachers at other schools within these circles for each reference school using the Common Core of Data (CCD) maintained by the NCES. Exposure to a state boundary within the local labor market is measured within these circles using count- and ratio-based metrics. 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 – say, for example, the circle with the 10-mile radius. School E is on the other side of a state line. Our ratio-based measure of boundary exposure in this example 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. Note that if School A were far from a state boundary, the value of the ratio would be zero because all nearby schools would be in the same state. We also estimate count models that include the numerator and

[4] Including charter schools, which mostly share pension and licensing policies with other public schools in the same state but sometimes do not, attenuates our primary findings slightly but they remain qualitatively similar.

denominator of Equation (2) separately, which are conceptually similar to the model used by Fitzpatrick and Lovenheim (2014).

The rationale underlying our approach is that for a given local labor market, frictions brought on by a state boundary will shrink the effective local labor pool. Our measures of boundary exposure are 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). Frictions may reduce total labor flows, affect which types of workers move, and/or affect workers' initial employment decisions. In education the latter issue could be particularly important because schools offer similar salaries and nonpecuniary job features matter more, making internal mobility more valuable (Greenberg and McCall, 1974). For example, 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 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. Principals may recognize this as well when forming their own preferences over jobs.

We use circles of 10- and 20-mile radii to define schools' local-area labor markets based on research showing that teachers have strong preferences for short commutes. For example, in their analysis of a large urban school district, Miller, Murnane and Willet (2008) find that the average teacher commutes just 7 miles. Similarly, Engel, Jacob and Curran (2014) find that teachers in Chicago 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), and Gershenson (2013) shows that substitute teachers are less likely to accept daily job offers that involve longer commutes.[5] These teacher-specific findings are consistent with recent research showing that workers generally have strong preferences for shorter commutes (Manning and Petrongolo, forthcoming). Our empirical analysis is consistent with evidence from these studies in that the student achievement effects of boundary exposure are driven by exposure within 10 miles.

In addition to our count- and ratio-based metrics that depend on teacher FTEs within 10 and 20 miles of a school, we also consider the robustness of our findings to a number of alternative boundary-exposure metrics. A simple modification is to include center-school FTE in our measures (e.g., in the denominator of Equation 2). We also replace the FTE-based metrics with metrics based on local-area student enrollment, and use metrics restricted to include only other schools in the reference school's local area with overlapping grades or similar student populations. In addition, we 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.

### *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]

---

[5] A related literature also shows that teachers exhibit preferences to work close to where they grew up (Boyd et al., 2005; Reininger, 2012).

[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.

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] Note that regardless of whether we have achievement data from a state, all schools in the lower-48 states and the District of Columbia are included when we code the geographic labor market areas for schools. Thus, the exclusion of a state from the main regression models owing to missing achievement data does not interfere with our ability to accurately code the out-of-state labor market shares for schools in neighboring states.

We also estimate the effect of boundary closeness on school proficiency rates. We normalize 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 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).

[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.

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 and the District of Columbia.[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. In some of our specifications below, we refer to this latter group as "intensely affected" by a state boundary. 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, 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 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 of test scores using our full sample of schools:

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

In Equation (3), $\psi_j$ is a state fixed effect and the covariates are the same as in Equation (1). The covariates 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 – i.e., those with 25 percent or more of local-area FTE in another state – and other schools are very small: using the 10-mile circles, they are 0.011 and 0.014 school-level standard deviations in math and reading, respectively, nominally favoring intensely-affected boundary schools. With the 20-mile circles the analogous gaps are -0.030 and -0.010 school-level standard deviations. None of the gaps are statistically significant.

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

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

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), $\rho_j$ 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 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. We show results from three variants of Equation (4) in Table 2. First, we code $\tilde{R}_{ij}$ as an indicator variable equal to one if the school is intensely affected by a state boundary using the 10-mile radius, per our definition above, and zero otherwise. Second, we estimate a similar model but define boundary exposure using the circles with 20-mile rather than 10-mile radii. In both cases we group 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 – and schools without any boundary exposure together and assign them a value of zero for the dependent variable. We also show results from an alternative coding where $\tilde{R}_{ij}$ captures the linear distance in miles to the nearest state boundary. In Appendix Table A3 we show results from several other versions of the selection model, which all corroborate the results in Table 2.

The top rows of Table 2 show full output for each model. We use the wild-cluster bootstrap to obtain confidence intervals for each coefficient from the selection regressions because our primary analytic sample includes just the 33 state clusters (Angrist and Pischke, 2008). 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]

[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

Table 2 shows results 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.

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, ranging from 0.45-0.92. These balancing results are achieved despite the fact that in most cases the confidence intervals for our estimates are not large, and shrink 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*

Tables 3 and 4 show the effects on math and reading achievement of exposure to a state boundary as estimated by two variants of Equation (1). First, in Table 3 we divide schools into three groups based on differential exposure to a state boundary: (a) "intensely affected" schools with 25 percent or more of local-area FTE on the other side of a state line, (b) "moderately affected" schools with more than zero but less than 25 percent of local-area FTE is on the other side of a state line, and (c) "unaffected" schools with no local-area FTE is on the other side of a state line.[10]

---

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] The distribution of the out-of-state FTE percentage (as shown in Equation 2) is shown in Appendix Figure A1 for schools in our analytic sample. Unfortunately, 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 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.

Unaffected schools are the omitted comparison group. We estimate models for math and reading achievement using the 10- and 20-mile radii to define local areas. Coefficients for the non-boundary covariates are suppressed in Table 3 (and subsequent tables) but can be found in Appendix A. Again, confidence intervals and statistical significance results are obtained via state-level wild-cluster bootstrapping.

Focusing first on the grade-8 math model, and the model that defines the local labor market using the 10-mile radius, 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. When we define the local labor market more broadly using the 20-mile measures, our results remain directionally similar but attenuate substantially. Consistent with the observational 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 models.

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 roughly 0.031 effect size in student-level standard deviations.[11]

[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).

Two aspects of the results in Table 3 suggest that the boundary effect is highly localized, an interpretation that is consistent with previous research on teacher commuting as described above. First is the attenuation of results as we expand the size of the local area around each school from 10 to 20 miles. Below, we further parse out the effect of boundary exposure as measured by the circles of 10- and 20-mile radii and confirm that exposure as measured by the 10-mile circles drives our findings. Table 3 also shows that schools where a smaller 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 confidence intervals, the weaker findings persist through many robustness and sensitivity analyses below, further implying that the effect of boundary exposure is highly localized.

Next, in Table 4, we estimate count-based models that are analogous to the models shown in Table 3. The count-based approach takes the ratio in Equation (2) and includes the numerator and denominator as separate terms (as in Fitzpatrick and Lovenheim, 2014). The models take the following form:

$$Y_{ij} = \gamma_0 + \mathbf{X_{ij}}\boldsymbol{\gamma_1} + \mathbf{FTE_{ij}}\boldsymbol{\gamma_2} + \mathbf{FTE_{ij}^{OS}}\boldsymbol{\gamma_3} + \pi_j + e_{ij} \quad (5)$$

The variable vectors $\mathbf{FTE_{ij}}$ and $\mathbf{FTE_{ij}^{OS}}$ include linear and quadratic terms that measure the number of FTE within the 10- or 20-mile radius, and within the radius and outside the state (OS), respectively. All other variables are the same as in Equation (1). The variables $\mathbf{FTE_{ij}}$ and $\mathbf{FTE_{ij}^{OS}}$ are coded so that they overlap; e.g., a school with 100 local area FTE, of which 25 are on the other side of a state line, would have values of $\mathbf{FTE_{ij}}$ and $\mathbf{FTE_{ij}^{OS}}$ of 100 and 25, respectively. Thus, $\boldsymbol{\gamma_3}$ can be interpreted as the effect of an increase in out-of-state FTE conditional on total local-area FTE. A benefit of the count-based models is that the association between having more FTE nearby and student achievement can be estimated at the same time as the boundary effect, but with the caveat that $\boldsymbol{\gamma}_2$ may not be causal because it is identified using variation between 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 4 are consistent with what we show in Table 3. First, note that the first and third columns of each panel document the relationship between local-area FTE and student achievement, omitting information about the location of 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 remain similar and out-of-state FTE has a negative effect on achievement conditional on total FTE. The results are again most pronounced using the 10-mile circles; they are attenuated but qualitatively similar using the 20-mile circles. To connect the estimates in Table 4 to the estimates in Table 3, note that 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 estimates in Table 4 for the math model with the 10-mile radius imply a test-score difference between school types of approximately -0.080 standard deviations, which is very close to the analogous estimate in Table 3.[13]

---

[12] At a minimum, $\mathbf{FTE_{ij}}$ can be viewed as serving the basic function of any other control variable in Equation (5). Unlike variation in $\mathbf{FTE_{ij}^{OS}}$, variation in $\mathbf{FTE_{ij}}$ does correlate significantly with many of the other covariates (results omitted for brevity), at least unconditionally, which raises concerns about its independent interpretation. Of course, we control for observed differences between schools with the ***X***-vector in our models, including population density (the most obvious potential confounder), which is helpful, but do not take a strong stand on whether unobserved correlates of $\mathbf{FTE_{ij}}$ contribute to the coefficient estimates.

[13] Appendix Table B6 shows an analogous version of Table 4 that reports on models that omit the quadratic FTE terms. The results are qualitatively similar.

*5.3 The Localness of Boundary Effects*

Table 5 reports on a sensitivity test regarding the local intensity of boundary effects suggested by our estimates in Tables 3 and 4. We divide the total labor market area within 20 miles of each school into two parts: (1) the part that is 0-10 miles from the school (i.e., within the circle of radius 10 miles), and (2) the part that is 11-20 miles from the school (i.e., within the 20 mile circle but outside of the 10-mile circle). We construct measures of boundary exposure analogous to what we use above based on local-area FTE within each distance range. These models allow us to test the effect of boundary exposure in the 11-20 mile range for each school conditional on exposure in the 0-10 mile range. Specifically, we estimate models of the following form:

$$Y_{ij} = \alpha_0 + \mathbf{X_{ij}}\boldsymbol{\alpha_1} + \mathbf{R_{ij}^{0\text{-}10}}\boldsymbol{\alpha_2} + \mathbf{R_{ij}^{11\text{-}20}}\boldsymbol{\alpha_3} + \theta_j + \xi_{ij} \tag{6}$$

Equation (6) is the same as Equation (1) except that that the 20-mile circle is divided into two parts within the equation. The parameter vector $\boldsymbol{\alpha_3}$ indicates how increased boundary exposure within 11-20 miles of the school, conditional on exposure within 10 miles, affects achievement.

Table 5 shows results for our ratio-based models, analogous in structure to the models in Table 3. It shows that conditional on how the local-area labor market is affected by a state boundary within 10 miles, differences in how the market is split 11-20 miles away has no discernable effect on achievement. This reinforces the point from above that the boundary effects are concentrated.

## 6. Robustness and Extensions

*6.1 Measurement and Models*

We examine the robustness of our findings to a variety of ways of measuring and modeling boundary exposure and comparing schools. Based on the preceding results showing that the 10-mile exposure measures are most informative, we restrict our attention to models that use these measures for the robustness and sensitivity tests. We relegate most of these analyses to Appendix B, where we consider: (a) measuring boundary exposure by 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 captured by the share of free/reduced-price lunch eligible students, (d) the use of an alternative, more-differentiated control group, (e) using exposure measures that include the school's own local-area FTE as part of the in-state FTE, (f) models that capture boundary exposure by linear and quadratic terms of the percentage of local-area FTE on the other side of a state line, (g) models that use the simple distance to a state boundary (linear and quadratic terms), and (h) regressions that are weighted by student enrollment and teacher FTEs. Summarizing the laundry-list of results in the appendix, our findings are qualitatively robust to the various modeling and measurement modifications.

### *6.2 Proficiency Rates*

In Table 6 we report results where we use school proficiency rates on state tests in place of 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 about the sample expansion. Table 6 presents results from proficiency-rate models that restrict the sample to include only states from Tables 3 and 4 (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. For ease of presentation we show results for the ratio-based models only using the 10-mile radius measure.

The findings in Table 6 are generally consistent with the results in Table 3. 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 6 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, 4, 7 and 8, along with the battery of tests in Appendix B).

### *6.2 Other Extensions*

Next we look for evidence of boundary effects in lower grades – grade-7, grade-5 and grade-3. There are two reasons to examine boundary effects in earlier grades. 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 labor frictions are responsible for our findings, it follows that boundary effects in lower grades should be smaller than in grade-8.[14]

Table 7 shows boundary effects on math test scores in grade-7, grade-5, and grade-3 using our ratio-based measures of boundary exposure and the 10-mile circles. The grade-7 point estimates are similar but slightly smaller than the garde-8 estimates in Table 3. The estimates for grades 5 and 3 are even smaller.[15] The pattern of 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

[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.

interventions have smaller effects on reading achievement than math achievement (e.g., Hanushek and Rivkin, 2010; Lefgren and Sims, 2012; Taylor and Tyler, 2012).

In another extension 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). The district-level measures capture the share of intensely and 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 outcome data are taken from the publicly-available Stanford Education Data Archive (SEDA; Reardon et al., 2016a) and derived from information about student performance 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, and Washington DC), which affords another opportunity to look at an expanded sample.

Table 8 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 and Washington DC is excluded in both models because SEDA data are unavailable). The results are consistent with our school-level findings and similar using the restricted and full samples. In the full math model with the district-aggregated data, the estimate in Table 8 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 ≈ 0.11) in the restricted 33-state sample.

Finally, we briefly discuss tests for heterogeneity in boundary effects. We had initially hoped heterogeneity analyses could be used to provide insights about the key drivers of boundary effects and possible moderators, but in practice our data structure and methods offer too little statistical power for heterogeneity analyses to be informative. In unreported results omitted for brevity, we tested for evidence of heterogeneous effects associated with differences in local-area population density, differences in average teacher wages between bordering states, and differences in state income-tax status. Across all of these dimensions we cannot statistically distinguish differential boundary effects; however, our standard errors are also too large to rule out meaningful heterogeneity. Thus, we do not draw strong inference from our tests for heterogeneous boundary effects.[16]

## 7. State Boundaries or District Boundaries?

Thus far we have established that schools with a larger 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 boundaries create labor frictions. 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.). It is of interest to understand if the achievement declines we see for schools near

[16] State level clustering has important power implications. Some progress may be possible if structural assumptions are imposed, but we do not pursue this approach here.

state boundaries are more than would be expected based on the coinciding incidence of district boundaries alone.

To test whether our findings indicate the presence of state-boundary effects above and beyond what would be expected owing to district boundaries alone, we add direct controls to our model for district-boundary exposure. Specifically, we estimate expanded models akin to what we show in Equations (1) and (3) as follows:

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

$$Y_{ij} = \tau_0 + \mathbf{X_{ij}}\boldsymbol{\tau_1} + \mathbf{FTE_{ij}}\boldsymbol{\tau_2} + \mathbf{FTE_{ij}^{OD}}\boldsymbol{\tau_3} + \mathbf{FTE_{ij}^{OS}}\boldsymbol{\tau_4} + \omega_j + e_{ij} \tag{8}$$

In equation (7), the variables in $\mathbf{R_{ij}^{OD}}$ measure the out-of-district FTE share and the variables in $\mathbf{R_{ij}^{OS}}$ continue to measure the out-of-state FTE share. Similarly, the terms $\mathbf{FTE_{ij}^{OD}}$ and $\mathbf{FTE_{ij}^{OS}}$ in Equation (8) include counts of out-of-district and out-of-state FTE. All other variables are as defined previously.

Note that 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 the out-of-district and out-of-state FTE controls are included simultaneously in the models, the coefficients on the out-of-district FTE variables ($\boldsymbol{\beta_2}$ or $\boldsymbol{\tau_3}$) are identified entirely from district boundaries within states. The coefficients on the out-of-state FTE variables ($\boldsymbol{\beta_3}$ or $\boldsymbol{\tau_4}$) capture the additional effect of out-of-state FTE conditional on the effect of out-of-district FTE exposure.

The results are shown in Tables 9 and 10 (corresponding to Tables 3 and 4 above). An important clarification for interpretation is that the omitted comparison group changes substantially in these new models. 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: for example, 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.[17]

The most important takeaway from Tables 9 and 10 is that district boundaries do not drive our findings for state boundaries. That said, the coefficients on the out-of-district FTE variables are also negative, which implies that district boundary frictions may also lower achievement. Combining the out-of-district and out-of-state FTE share coefficients in the math model with the 10-mile radius in Table 9, and taking the estimate for out-of-district FTE at face value, implies 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 – the implied effect is a relative reduction of test scores in math of -0.1571 (-0.1005 + (-0.0566)) school-level standard deviations, or roughly -0.05 student-level standard deviations.[18]

## 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 negative effects that are only sometimes statistically significant. Our estimates can be converted into student-level standard deviations, which are more commonly used in education research, by multiplying them by roughly one-third (Bhatt and Koedel, 2012; Burgess, Wilson and Worth, 2013). Although the boundary effects are small on a per-student basis, they are spread across a very large population: based on the Common Core of Data, we

[17] 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 in the ratio-based models (Equation 7). 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 ratios for presentational convenience in Equation (7).
[18] 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.

estimate that roughly 670,000 students are enrolled in middle schools nationally that are coded as "intensely affected" by a state boundary in our study.

Labor frictions at state boundaries are a plausible explanation for our findings. A large literature in economics documents the adverse effect of labor frictions 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. We also note that while we put forth state-specific pension and licensing policies as the most likely factors driving frictions in teacher labor markets near state boundaries, other state policies may also create frictions. Possibilities include aforementioned differences across states in teacher salaries and tax policies, among others. We attempted to examine heterogeneity in boundary effects along these dimensions, but our heterogeneity analyses are underpowered. 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]

A large literature on teacher quality shows that teachers are important inputs into the educational production function and that teacher effectiveness is influenced by the match with the school (Jackson, 2013). Research has focused primarily on estimating achievement effects of exposure to a more effective teacher in a single year, and as such it is difficult to directly connect previous findings to our results, which reflect the cumulative effect of boundary exposure through

[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.

grade-8. That said, with some assumptions we can perform a back-of-the-envelope calculation to approximate the implied effect of boundary exposure on the quality of instruction during the grades covered by our analysis. Specifically, we assume that the standard deviation of teacher effectiveness is 0.14 student standard deviations each year in grades K-8 (e.g., per Chetty, Friedman and Rockoff, 2014), boundaries have the same effect on the labor market in all grades K-8, and we allow for the decay of teacher effects over time based on available estimates (Chetty, Friedman and Rockoff, 2014; Jacob, Lefgren and Sims, 2010). Under these conditions, our estimate of the cumulative boundary effect in math of 0.031 student standard deviations by grade-8 implies that intense exposure to a state boundary lowers teacher effectiveness by about 0.06-0.12 standard deviations of the teacher distribution.[20]



[20] This range of estimates depends in part on an assumption about student mobility between boundary and non-boundary schools, which affects the number of years of boundary exposure for observed grade-8 students in the treatment and control conditions. The lower end of the range reported in the text assumes students do not switch between boundary and non-boundary schools at all in grades K-8.

Acknowledgments
Kim, Ni and Podgursky are in the Department of Economics, and Koedel is in the Department of Economics and Truman School of Public Affairs, at the University of Missouri, Columbia. The authors thank Maurice Li for research assistance and Dan Goldhaber, Peter Mueser and Eric Parsons for useful comments. 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 #R305C120008 to American Institutes for Research from the Institute of Education Sciences, U.S. Department of Education. The views expressed here are those of the authors and should not be attributed to their institutions, data providers, or the funders. Any and all errors are attributable to the authors.

References

Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

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), 367-77.

Bandeira de Mello, Victor. 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, Victor, 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, Hamilton Lankford, Susanna Loeb and James Wyckoff. 2005. The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools. *Journal of Policy Analysis and Management* 24(1), 113–32.

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.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90(3), 414–27.

Cannata, Marissa. 2010. Understanding the Teacher Job Search Process: Espoused Preferences and Preferences in Use. *Teachers College Record* 112(12), 2889-934.

Chetty, Raj, John N. Friedman and Jonah E. Rockoff. 2014. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review* 104(9), 2593-632.

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), 1191-230.

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.

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), 269-79.

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

Gershenson, Seth. 2013. The Causal Effect of Commute Time on Labor Supply: Evidence from a Natural Experiment Involving Substitute Teachers. *Transportation Research Part A: Policy and Practice* 54, 127–40.

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), 421-31.

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), 1100–37.

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), 1096–116.

Jacob, Brian A., Lars Lefgren and David P. Sims. 2010. The Persistence of Teacher-Induced Learning Gains. *Journal of Human Resources* 45(4), 915-943.

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 Handbook of Economics of Education Vol. 5, eds. Eric A. Hanushek, Stephen Machin and Ludger Woessman: 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.

Manning, Alan and Barbara Petrongolo (forthcoming). How Local Are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*.

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. An Examination of the Movement of Educators within and across three Midwest Region 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, Benjamin Shear, Kenneth Shores and Erin Fahle. 2016a. *Stanford Education Data Archive*. http://purl.stanford.edu/db586ns4974.

Reardon, Sean, Benjamin Shear, Katherine Castellano and Andrew Ho. 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.

Reininger, Michelle. 2012. Hometown Disadvantage? It Depends on Where You're From: Teachers' Location Preferences and the Implications for Staffing Schools. *Educational Evaluation and Policy Analysis* 34(2), 127–45.

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

Solon, Gary, Steven J. Haider and Jeffrey M. Wooldridge. 2015. What Are We Weighting For? *Journal of Human Resources* 50(2), 301–16.

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.

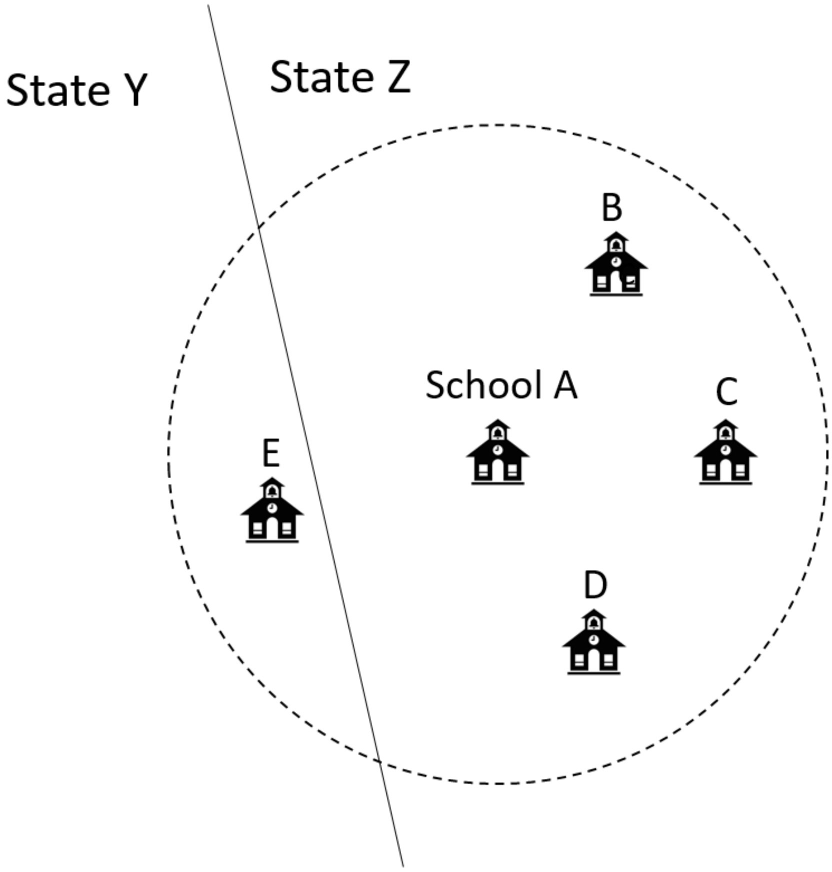
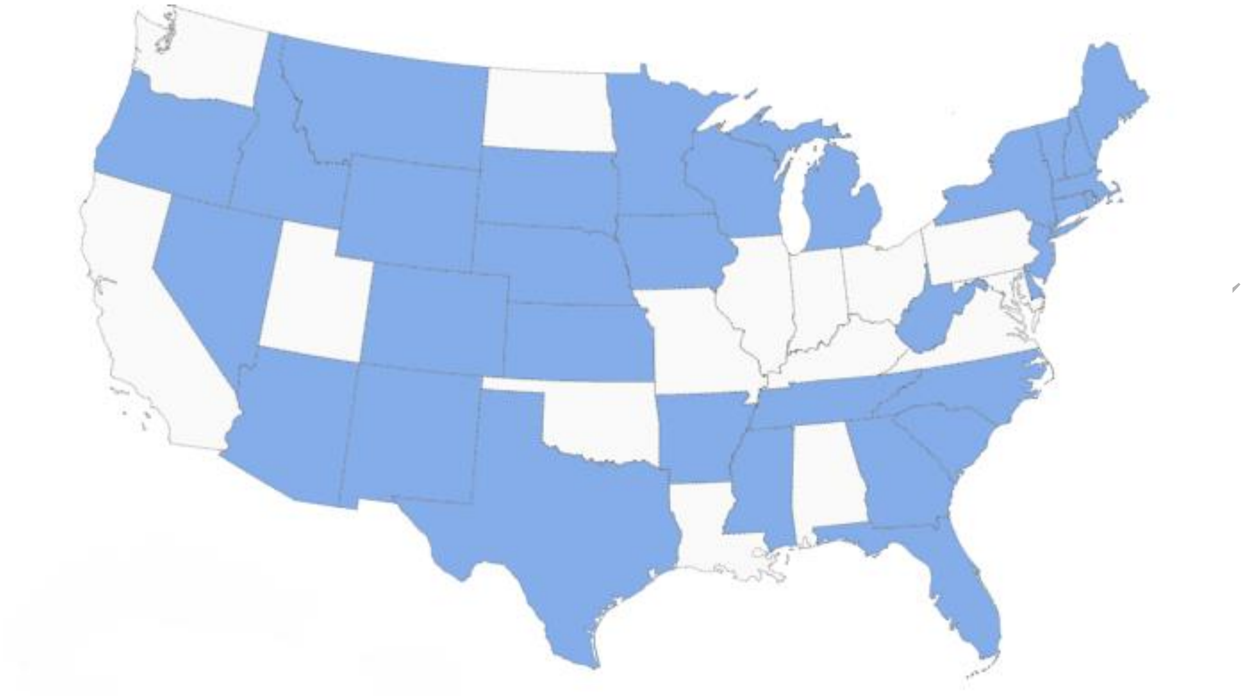


Figure 2: 33 States with Grade-8 Math Scaled Scores Included in the Primary Analytic Sample.



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 |
| *Labor Market Bifurcation (10-mile Circle)* | | | | |
| Out-of-state Labor Market Percent ≥ 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) |
|---|---|---|---|---|---|---|
| | 10-mile radius | | 20-mile radius | | | |
| VARIABLES | 1(Out-of-State Labor Market Percent ≥ 25) | | | | Distance to a state border | |
| % Free Lunch Status | 0.0003 | 0.0002 | 0.0002 | 0.0000 | -0.1476 | -0.1475 |
| | [-0.000,0.001] | [-0.000,0.001] | [-0.001,0.001] | [-0.001,0.001] | [-0.490,0.165] | [-0.420,0.131] |
| % Reduced Lunch Status | -0.0004 | 0.0002 | -0.0008 | 0.0002 | -0.4580 | -0.5440 |
| | [-0.001,0.001] | [-0.000,0.001] | [-0.003,0.001] | [-0.001,0001] | [-1.270,0.373] | [-1.199,0.122] |
| % Asian | 0.0016 | 0.0007 | -0.0002 | -0.0020 | -0.9283 | -0.4320 |
| | [-0.002,0.005] | [-0.002,0.003] | [-0.003,0.002] | [-0.004,0.000] | [-1.938,0.112] | [-1.009,0.172] |
| % Hispanic | -0.0004 | 0.0001 | -0.0007 | 0.0002 | 1.3741* | 0.9889 |
| | [-0.001,0.000] | [-0.000,0.001] | [-0.002,0.001] | [-0.001,0.001] | [0.649,2.120] | [0.214,1.729] |
| % Black | -0.0005 | -0.0004 | -0.0002 | -0.0002 | 0.0076 | -0.0253 |
| | [-0.001,0.000] | [-0.001,0.000] | [-0.002,0.001] | [-0.000,0.000] | [-0.268,0.308] | [-0.262,0.208] |
| % American Indian | -0.0003 | 0.0001 | -0.0002 | 0.0006 | 0.2680 | 0.1506 |
| | [-0.001,0.000] | [-0.001,0.001] | [-0.002,0.001] | [-0.000,0.002] | [-0.067,0.623] | [-0.196,0.496] |
| % Pacific Islander | 0.0235 | 0.0160 | 0.0284 | 0.0125 | -6.5280 | -0.8220 |
| | [-0.001,0.047] | [-0.003,0.035] | [0.000,0.056] | [-0.005,0.031] | [-14.65,1.373] | [-3.553,2.010] |
| % Two or more race | -0.0009 | 0.0006 | -0.0008 | 0.0009 | -0.3110 | -0.9368 |
| | [-0.004,0.003] | [-0.001,0.003] | [-0.001,0.007] | [-0.003,0.005] | [-2.063,1.546] | [-2.358,0.447] |
| Log of Total Enrollment | 0.0089 | 0.0061 | 0.0128 | 0.0007 | -7.1819 | -3.7988 |
| | [-0.002,0.019] | [-0.001,0.014] | [-0.009,0.033] | [-0.012,0.014] | [-17.33,3.639] | [-8.725,1.361] |
| Log of Total District Enrollment | -0.0115 | -0.0134 | -0.0132 | -0.0105* | 7.4215 | 3.4269 |
| | [-0.027,0.003] | [-0.027,0.000] | [-0.032,0.005] | [-0.022,0.002] | [-6.122,20.42] | [-1.213,8.076] |
| % English Language Learners | -0.0025 | -0.0014 | -0.0048 | -0.0026 | -0.1357 | -0.0755 |
| | [-0.005,0.000] | [-0.003,0.000] | [-0.009,-0.000] | [-0.005,-0.000] | [-0.840,0.470] | [-0.769,0.573] |
| Log of Total Revenue per pupil | 0.0554 | -0.0009 | 0.1782 | 0.0596 | -42.9486** | -6.3986 |
| | [0.006,0.102] | [-0.042,0.040] | [0.050,0.302] | [-0.005,0.127] | [-73.10,-11.53] | [-36.11,23.71] |
| Log of Local Revenue per pupil | -0.0181*** | -0.0089** | -0.0438** | -0.0257 | 15.8403 | 0.6036 |
| | [-0.031,-0.006] | [-0.018,-0.001] | [-0.075,-0.012] | [-0.054,0.003] | [-4.763,35.77] | [-21.48,22.59] |
| Log Median Household Income | 0.0577 | 0.0222 | 0.1336* | 0.0509 | -30.4040** | -16.4235 |
| | [0.010,0.104] | [-0.004,0.048] | [0.033,0.229] | [-0.008,0.112] | [-45.28,-15.22] | [-34.59,2.082] |
| % Low Education | 0.0016 | 0.0005 | 0.0030 | 0.0007 | -0.6139** | -0.4469 |
| | [0.000,0.003] | [-0.000,0.001] | [0.001,0.005] | [-0.000,0.002] | [-1.026,-0.241] | [-0.856,-0.043] |
| Population Density/1000 | 0.0096 | 0.0064 | 0.0225 | 0.0147 | -2.4122 | -1.3667 |
| | [0.001,0.018] | [-0.001,0.014] | [0.005,0.040] | [0.001,0.028] | [-4.871,0.107] | [-3.838,1.134] |
| 1(Suburb) | 0.0004 | -0.0100 | 0.0311 | 0.0063 | 14.6649 | 5.9899 |
| | [-0.016,0.017] | [-0.031,0.011] | [-0.018,0.080] | [-0.029,0.041] | [-9.698,38.24] | [-6.346,18.41] |
| 1(Town) | -0.0214 | -0.0287** | 0.0026 | -0.0066 | 8.1953 | 6.0233 |
| | [-0.048,0.004] | [-0.056,-0.002] | [-0.052,0.054] | [-0.051,0.040] | [-13.83,30.01] | [-6.108,18.10] |
| 1(Rural) | -0.0357** | -0.0342** | -0.0174 | -0.0091 | 20.6961 | 12.4977 |
| | [-0.069,-0.002] | [-0.070,-0.001] | [-0.071,0.041] | [-0.054,0.036] | [-5.325,46.28] | [-0.191,25.19] |
| Constant | -0.9563 | -0.1058 | -2.7289* | -0.8328 | 648.8718*** | 313.9632* |
| | [-1.653,-0.251] | [-0.597,0.397] | [-4.300,-1.194] | [-1.964,0.381] | [440.4,872.0] | [44.36,572.1] |
| State Fixed Effects | | X | | X | | X |
| R-squared | 0.0540 | 0.1164 | 0.1204 | 0.2510 | 0.2650 | 0.4946 |
| Observations (schools) | 11,686 | 11,686 | 11,686 | 11,686 | 11,686 | 11,686 |
| Joint P-Value | | 0.45 | | 0.92 | | 0.91 |

Notes: This table shows variants of the selection equation in the main text where we adjust the dependent variable. The first four columns predict an indicator for being an intensely-affected boundary school using available covariates. Columns 1 and 2 use 10-mile radius in defining local labor market; columns 3 and 4 use 20-mile radius in defining local labor market. Columns 5 and 6 are from a model where the dependent variable is the distance to a closest state border (linear). Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008). *** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimated Boundary Effects on Grade 8 Scaled Scores, Ratio Model.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | 10-mile radius | | 20-mile radius | |
| | Grade 8 Scaled Score | | | |
| VARIABLES | Math | Reading | Math | Reading |
| Out-of-State Percent ≥ 25 | -0.0942** | -0.0537* | -0.0521* | -0.0112 |
| | [-0.164, -0.026] | [-0.108, -0.002] | [-0.104,-0.001] | [-0.055,0.033] |
| 0< Out-of-State Percent <25 | -0.0095 | 0.0082 | 0.0287 | 0.0167 |
| | [-0.070, 0.048] | [-0.028, 0.047] | [-0.039,0.093] | [-0.027,0.058] |
| Covariates | X | X | X | X |
| State Fixed Effects | X | X | X | X |
| R-squared | 0.4773 | 0.5896 | 0.4773 | 0.5895 |
| Observations (schools) | 11,686 | 13,286 | 11,686 | 13,286 |

Notes: Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Estimated Boundary Effects on Grade 8 Scaled Scores, Count Model.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | 10-mile radius | | | | 20-mile radius | | | |
| VARIABLES | Math | Math | Reading | Reading | Math | Math | Reading | Reading |
| Total FTE/1000 | 0.0129 | 0.0142 | 0.0214** | 0.0252** | 0.0043 | 0.0057 | 0.0067** | 0.0073*** |
| | [0.001,0.024] | [-0.001,0.029] | [0.001,0.033] | [0.009,0.041] | [-0.001,0.010] | [-0.000,0.012] | [0.002,0.011] | [0.002,0.013] |
| (Total FTE/1000)$^2$ | -0.0002 | 0.0001 | -0.0003** | -0.0003 | -0.0000 | -0.0000 | -0.000 | -0.0000 |
| | [-0.000,0.000] | [-0.001,0.001] | [-0.001,-0.000] | [-0.001,0.000] | [-0.000,0.000] | [-0.000,0.000] | [-0.000,0.000] | [-0.000,0.000] |
| Out-of-State FTE/1000 | | -0.0389*** | | -0.0362*** | | -0.0111*** | | -0.0118*** |
| | | [-0.055,-0.025] | | [-0.053,-0.020] | | [-0.017,-0.005] | | [-0.019,-0.005] |
| (Out-of-State FTE/1000)$^2$ | | 0.0007 | | 0.0010 | | 0.0001 | | 0.0001 |
| | | [-0.000,0.002] | | [0.000,0.002] | | [-0.000,0.000] | | [-0.000,0.000] |
| Covariates | X | X | X | X | X | X | X | X |
| State Fixed Effects | X | X | X | X | X | X | X | X |
| R-squared | 0.4835 | 0.4841 | 0.5959 | 0.5965 | 0.4797 | 0.4799 | 0.5933 | 0.5937 |
| Observations (schools) | 11,686 | 11,686 | 13,286 | 13,286 | 11,686 | 11,686 | 13,286 | 13,286 |

Notes: Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Estimated Boundary Effects on Grade 8 Scaled Scores Adding 11-20 mile Labor Market Variables, Ratio Model.

| | (1) | (2) |
|---|---|---|
| | Grade 8 Scaled Score | |
| VARIABLES | Math | Reading |
| 10 mile Out-of-State Labor Market Percent ≥ 25 | -0.0856** | -0.0469 |
| | [-0.170,-0.006] | [-0.115,0.023] |
| 0 < 10 mile Out-of-State Labor Market Percent < 25 | -0.0050 | 0.0119 |
| | [-0.074,0.062] | [-0.042,0.066] |
| 11-20 mile Out-of-State Labor Market Percent ≥ 25 | -0.0107 | -0.0079 |
| | [-0.060,0.038] | [-0.071,0.051] |
| 0 < 11-20 mile Out-of-State Labor Market Percent < 25 | 0.0329 | 0.0297 |
| | [-0.042,0.104] | [-0.021,0.077] |
| Covariates | X | X |
| State Fixed Effects | X | X |
| R-squared | 0.4775 | 0.5897 |
| Observations (schools) | 11,686 | 13,286 |

Notes: Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Estimated Boundary Effects on Grade 8 Proficiency Rate, Ratio Model.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Scaled Score Sample | | Extended Sample | |
| VARIABLES | 10-mile radius | | 10-mile radius | |
| | Math | Reading | Math | Reading |
| Out-of-State Percent ≥ 25 | -0.0816** | -0.0329 | -0.0772** | -0.0319 |
| | [-0.145,-0.020] | [-0.089,0.022] | [-0.130,-0.022] | [-0.102,0.038] |
| 0<Out-of-State Percent<25 | -0.0082 | 0.0412 | 0.0160 | 0.0761** |
| | [-0.085,0.066] | [-0.015,0.096] | [-0.067,0.090] | [0.019,0.133] |
| 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 in Table 3, are because scale scores and proficiency rates are not both available for all schools. Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Robustness Results Estimated Using Grade-7, Grade-5 and Grade-3 Scaled Scores; Ratio Model With 10-mile Radius.

| | (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 |
| | [-0.140,-0.005] | [-0.112,0.021] | [-0.123,0.031] | [-0.097,0.017] | [-0.117,0.030] | [-0.070,0.079] |
| 0<Out-of-State Percent<25 | 0.0073 | 0.0369 | -0.0241 | 0.0065 | 0.0112 | -0.0132 |
| | [-0.065,0.085] | [-0.012,0.085] | [-0.133,0.078] | [-0.063,0.070] | [-0.060,0.084] | [-0.078,0.056] |
| 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 | 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; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 8: 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** [-0.074,-0.009] | -0.0192 [-0.041,0.007] | -0.0430** [-0.074,-0.011] | -0.0260** [-0.045,-0.006] |
| Share of "Moderately Affected" Boundary Schools | -0.0062 [-0.045,0.029] | 0.0100 [-0.020,0.041] | 0.0005 [-0.030,0.032] | 0.0040 [-0.021,0.027] |
| 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; Washington DC is excluded in all columns. SEDA does not provide data from Washington DC or grade-8 math data from California, for the same reason that we do not include these data (see textStandard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Estimated Boundary Effects on Grade 8 Scaled Scores with Out-of-District Variables, Ratio Model Using 10-mile Radius.

| | (1) | (2) |
|---|---|---|
| | Grade 8 Scaled Score | |
| VARIABLES | Math | Reading |
| Out-of-State Labor Market Percent ≥ 25 | -0.1005** | -0.0578** |
| | [-0.171,-0.030] | [-0.114,-0.005] |
| 0 < Out-of-State Labor Market Percent < 25 | -0.0101 | 0.0080 |
| | [-0.070,0.047] | [-0.028,0.047] |
| Out-of-District & In-State Labor Market Percent ≥ 25 | -0.0566*** | -0.0539*** |
| | [-0.095,-0.023] | [-0.083,-0.025] |
| 0 < Out-of-District & In-State Labor Market Percent < 25 | -0.0423 | -0.0564 |
| | [-0.110,0.022] | [-0.117,0.004] |
| Covariates | X | X |
| State Fixed Effects | X | X |
| R-squared | 0.4777 | 0.5899 |
| Observations (schools) | 11,686 | 13,286 |

Notes: Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Estimated Boundary Effects on Grade 8 Scaled Scores with Out-of-District Variables, Count Model Using 10-mile Radius.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Grade 8 Scaled Score | | | |
| VARIABLES | Math | | Reading | |
| Total FTE/1000 | 0.0129 | 0.0193 | 0.0214** | 0.0567 |
| | [0.001,0.024] | [-0.006,0.047] | [0.010,0.033] | [0.001,0.114] |
| $(\text{Total FTE}/1000)^2$ | -0.0002 | 0.0009*** | -0.0003** | -0.0011 |
| | [-0.000,0.000] | [0.000,0.002] | [-0.001,-0.000] | [-0.003,0.001] |
| Out-of-State FTE/1000 | | -0.0635*** | | -0.0520 |
| | | [-0.089,-0.039] | | [-0.100,-0.006] |
| $(\text{Out-of-State FTE}/1000)^2$ | | -0.0004 | | 0.0020 |
| | | [-0.001,0.001] | | [-0.000,0.004] |
| Out-of-District FTE/1000 | | -0.0056 | | -0.0425 |
| | | [-0.035,0.023] | | [-0.095,0.008] |
| $(\text{Out-of-District FTE}/1000)^2$ | | -0.0012** | | 0.0014 |
| | | [-0.002,-0.000] | | [-0.001,0.004] |
| Covariates | X | X | X | X |
| State Fixed Effects | X | X | X | X |
| R-squared | 0.4835 | 0.4975 | 0.5959 | 0.6094 |
| Observations (schools) | 11,686 | 11,686 | 13,286 | 13,286 |

Notes: Standard errors are clustered at the state level; statistical significance and confidence intervals are based on the wild cluster bootstrap-t procedure as described by Cameron et al. (2008).
*** p<0.01, ** p<0.05, * p<0.1