Online Appendix

"The Effects of Occupational Licensing Stringency on the Market for Public School Teachers"

Bradley J. Larsen Ziao Ju Adam Kapor Chuan Yu[30]

A Additional Data and Estimation Details

A.1 Data Construction, Sample Restrictions, and Weights

This section provides additional details about the data construction. We obtain data on individual teachers from the SASS surveys. We limit our sample to teachers from school districts with at least 50 students enrolled. Given that we have 17 years and 51 states (for our purposes, Washington, D.C. is also treated as a state), we would have 867 state-year cells constructed from 26,320 teachers. We restrict our sample to state-year cells with at least six teachers to satisfy NCES minimum cell-size requirements (described in more detail below), reducing the number of individual teachers to 26,280 and the number of state-year cells to 857. Note that, in our paper, all raw data sample sizes from NCES are rounded to the nearest 10 to comply with reporting requirements.

College selectivity is measured by the average SAT score of entering freshmen from a 1983 survey conducted by the Higher Education Research Institute (see Astin et al. 1983 and Angrist and Guryan 2008).[31] The state-year moments of teachers' college selectivity (as well as the overall mean and standard deviation used for standardization) are computed using individual-level SASS sampling weights, and this is true of all state-year controls constructed from SASS data. We aggregate to the group (state-year) level because our policy variation (licensing stringency) exists at the state-year level and our empirical analysis in Section 4

[30]Larsen: Washington University in St. Louis and NBER; blarsen@wustl.edu. Ju: Analysis Group; ziao.ju@analysisgroup.com. Kapor: Princeton University and NBER; akapor@princeton.edu. Yu: Department of Economics, University of North Carolina; chuanyu@unc.edu.

[31]The average SAT score of an institution is highly correlated over time. To verify this, we obtained 2007 institutional-level data from `collegescorecard.ed.gov` (U.S. Department of Education, 2020a). For the set of institutions appearing in both the 1983 and 2007 data, the correlation of average SAT scores across the two years is 0.81.

relies on grouped quantile regression. Additionally, aggregating (as well as standardizing the college selectivity measure) allows our analysis to rely on data that can be released to other researchers while satisfying NCES reporting requirements.

All earnings amounts in the paper (e.g., Table 1) are expressed in 2007 dollars. One earnings variable in Table 1.A is the median log earnings of *private* school teachers in a state-year cell. This variable comes from separate SASS surveys of private school teachers, available in the same years as the SASS public school surveys. These data are more sparse than the public school teacher sample, and fewer state-year cells constructed from these data meet our minimum cell-size requirement of six. We set this variable to zero for any state-year cells with fewer than six observations and construct a dummy indicator for these cells, denoted "Pri. teacher earnings exists" in Table 1.A.

In the heterogeneity analysis of Section 4.4, splitting the data into high- versus low-poverty (or high- versus low-minority) observations should yield two observations per state-year pair instead of one, meaning the number of observations in this analysis should be double the number in Table 3. In practice, however, we have fewer observations for several reasons. First, the percent of students qualifying for free lunch is constant within some state-year cells, which leads to only a single observation in these cells in the aggregated state-year-poverty-status data. Another reason is that NCES does not allow reporting data from cells with fewer than three observations. We therefore set the minimum cell size at the state-year level to six, which implies that when the data are split into high- versus low-poverty observations, the cell size typically shrinks to three, except in cases where more than half of the observations in a state-year cell have the same value for the percent of students qualifying for free lunch. In such cases, to comply with reporting requirements, we drop any state-year-poverty-status cell with fewer than three observations. The same reasoning applies when we consider high- versus low-minority observations.

All regressions that use SASS-based outcomes are weighted regressions using SASS sampling weights aggregated to the state-year level. In the teacher quantity analysis in Section 4.5, the outcomes are based on CCD or IPEDS data rather than SASS, so we do not use SASS sampling weights in these regressions and instead rely only on non-SASS controls. This yields 867 state-year cells rather than 857, as the NCES minimum cell-size requirement no longer applies.

Appendix Table A.7 lists the sources for each of the 37 teacher licensing requirements for each year. The table shows that data on some of the 37 certification requirements are missing in certain years; additional requirements beyond the 37 we focus on are also recorded in some manuals but are missing for more years. For any of the 37 requirements with missing data in 1991 (which applies to only two of the 37 requirements), we fill in these missing values using the next available year. We fill in any other missing requirement-year cells with the most recent available information. When data are available for a given requirement in a given year, they are available for *all* states. Appendix Table A.7 also reports the mean and standard deviation of each requirement, computed by first taking the mean and standard deviation within each state over time and then averaging across states.

The Bartik labor demand variable is constructed as follows. For state s and year t, this variable is defined as

$$B_{s,d(t)} \equiv \frac{1}{100,000} \sum_{k} L_{sk} g_{kd},$$

where L_{sk} is industry k's share of employment in state s in 1990 and g_{kd} is the national wage growth rate of industry k from decade $d-1$ to decade d. The notation $d(t)$ denotes the decade surrounding year t. Therefore, this measure varies by decade within a given state: for years 1991–2000, the variable takes the 1991–2000 decade value, and for years 2001–2008, it takes the 2001–2008 decade value. Data used to construct this variable come from the 1990 and 2000 U.S. Census IPUMS data and the 2006–2008 American Community Survey (Ruggles et al. (2018)).

The Klarner (2013a,b) data are coded as follows. For the executive branch, 0 = Republican, 1 = Democratic, and 0.5 = Independent or other. For the legislative branch, 0 = Republican control of both houses; 0.25 = Republican control of one house and split control of the other; 0.5 = Democratic control of one house and Republican control of the other; 0.75 = Democratic control of one house and split control of the other; and 1 = Democratic control of both houses. The Klarner (2013a,b) data, as well as the collective bargaining data, do not contain information on Washington, D.C.; we manually collect these variables for D.C., primarily from Wikipedia, treating the mayor as the executive branch and the Council of the District of Columbia as the legislative branch.

A.2 Additional Estimation Details

Grouped Quantile Regression (GQR). Our main estimator is a special case of the grouped quantile regression estimator (Chetverikov et al. 2016), an alternative to quantile regressions for settings with group-level treatment and micro data on outcomes within a group. In our paper, a group is a state-year cell, the treatment is state-year licensing stringency, and the micro data correspond to individual teachers surveyed in each cell. This estimator offers a number of advantages over traditional quantile regressions. In particular, it is computationally simple to estimate (via ordinary least squares) even with a large number of fixed effects and is robust to measurement error in computing quantiles (unlike standard quantile regressions; see Chetverikov et al. 2016 for details).

Like the unconditional quantile regression (UQR) approach of Firpo et al. (2009), the special case of the GQR estimator we use estimates an *unconditional* effect (i.e., an effect that does not depend on the level of the right-hand-side variable of interest). A distinction between GQR and UQR is that the former estimates an effect on the outcome distribution within a group, whereas the latter estimates an effect on the distribution in the full sample.

Our minimum cell size (the number of observations per group) is six in our main analysis, so that the 10th percentile in the smallest cells is equivalent to the 20th percentile and the 90th percentile is equivalent to the 80th percentile. Similarly, in our heterogeneity analysis of Section 4.4, our minimum cell size is three, and in these smallest cells, several percentiles of the distribution coincide. Chetverikov et al. (2016) derived a mild growth condition on the minimum cell size as the number of cells increases such that this measurement error can also be ignored when computing standard errors.

R^2 with Different State-Year Controls. In Table 3, we add different sets of state-year control variables. Across these specifications, the R^2 changes very little as we move from a specification with only state and year fixed effects (column 1) to our preferred specification with all controls (column 7). This stability suggests that factors beyond state and year fixed effects — including the various education policy, labor market, political, and demographic controls we include — explain little of the remaining variation in the outcome. Licensing stringency, on the other hand, has a consistent and significant estimated effect across these specifications. These R^2 results are reported in Appendix Table A.4.

Replication of Angrist and Guryan (2008). Our replication of Angrist and Guryan (2008) (AG) in Table 4 does not perfectly align with AG for several reasons. First, AG use SASS data from the 1987, 1993, and 1999 surveys, and in their analysis of first-year teachers, they include only teachers who began teaching in the survey year. In contrast, our construction also uses, for example, second-year teachers in the 1993 survey and links the survey data for those teachers to state licensing requirements in 1992, the year they were initially licensed.[32] Second, AG run regressions where the left-hand-side variable is the average of teachers' college selectivity in a given school district and year rather than at the state-year level. AG use state-year certification exam requirements as the right-hand-side variable. Third, AG include a subset of our state-year controls — urbanicity fixed effects and a quadratic in unemployment — and control for enrollment and the fraction of minority students at the district rather than the state level ("AG Controls" in Table 4 refers to these controls but with the latter two measured at the state rather than district level). Finally, AG do not standardize the college selectivity measure. Column 1 of Table 4.A most closely corresponds to the specification reported in AG Table 5, column 4.

B Teacher Attrition

As outlined in Section 4, our estimation approach relies on a state-by-year panel from 1991–2007 constructed using four SASS survey waves. We group together teachers who started teaching in the same year and the same state, and who therefore faced the same licensing requirements when initially licensed. For example, observations for the year 1993 in our state-by-year panel correspond to teachers in their first year of teaching when observed in the 1993 SASS survey; observations for 1999, 2003, and 2007 similarly contain first-year teachers. Observations for years 1992, 1998, 2002, and 2006 correspond to teachers observed in a SASS survey during their second year of teaching.

Constructing the dataset in this way could potentially bias estimates of the treatment effect of licensing stringency if teachers with different levels of college selectivity differ in their likelihood of exiting the profession. In that case, changes in the distribution of college selectivity could partly reflect teacher attrition rather than the direct effect of licensing policy. We provide a simple theoretical illustration of this potential bias and then show that

[32] Note that AG use data from 1990 for some analyses, but not for their analysis of college selectivity, as information on teachers' undergraduate institutions is missing in that year.

it does not appear to be important in our empirical analysis.

For simplicity, consider two states, A and B, where state A has lower licensing stringency than state B. We compare the difference in college selectivity between teacher cohorts in A and B when they are observed in their first year of teaching versus their second year, after some teachers may have exited the profession. Suppose teachers' college selectivity is uniformly distributed on $[q_L^s, q_H^s]$ when a cohort begins teaching in state $s \in \{A, B\}$.

Assume the attrition rate, defined as the probability that a teacher exits the profession in a given year, is equal to $\underline{a}$ for teachers with college selectivity $q \leq \phi$ and $\overline{a}$ for teachers with $q > \phi$, where $0 \leq \underline{a} \leq \overline{a} < 1$ and $\max\{q_L^A, q_L^B\} \leq \phi \leq \min\{q_H^A, q_H^B\}$.[33] Thus, teachers with college selectivity below ϕ exit with a weakly lower probability, implying that attrition is weakly increasing in college selectivity.[34]

Let $q_{\tau,y}^s$ denote the τth percentile of college selectivity among the cohort in state s in year y of teaching. In year 1, this is given by

$$q_{\tau,1}^s \equiv q_L^s + \frac{\tau}{100}(q_H^s - q_L^s).$$

After one year, attrition may change the distribution. Consider first a relatively low percentile τ such that $\max\{q_{\tau,1}^A, q_{\tau,1}^B\} \leq \phi$, for example, the 10th percentile. In year 2, a fraction

$$(1 - \underline{a})(\phi - q_L^s) + (1 - \overline{a})(q_H^s - \phi)$$

of teachers remain in the profession. The τth percentile among *second-year* teachers in state s is therefore

$$q_{\tau,2}^s \equiv q_L^s + \frac{\tau}{100}\left(\frac{1 - \overline{a}}{1 - \underline{a}}(q_H^s - \phi) + \phi - q_L^s\right).$$

The difference between the τth percentile among first-year teachers in the two states is

$$\Delta_{\tau,1} \equiv q_{\tau,1}^B - q_{\tau,1}^A = (q_L^B - q_L^A) + \frac{\tau}{100}(q_H^B - q_H^A - q_L^B + q_L^A).$$

[33] We assume the attrition function is the same in both states.

[34] Lankford et al. (2002) provided evidence that teacher attrition may be higher among teachers with greater outside earnings opportunities, and Wiswall (2007) incorporates this feature into a model of teacher labor supply. Other studies similarly suggest that the pursuit of higher paying alternatives is an important reason teachers leave the profession (Kirby et al. 1999; Prince 2003; Hanushek et al. 2004).

For second-year teachers, the difference is

$$\Delta_{\tau,2} \equiv q^B_{\tau,2} - q^A_{\tau,2} = (q^B_L - q^A_L) + \frac{\tau}{100}\left(\frac{1-\overline{a}}{1-\underline{a}}(q^B_H - q^A_H) - q^B_L + q^A_L\right).$$

The object $\Delta_{\tau,1}$ measures the difference in the τth percentile between states A and B that arises solely from the difference in licensing stringency faced by the cohort when entering the profession, whereas $\Delta_{\tau,2}$ captures differences that reflect both the direct effect of licensing stringency and the effect of attrition. The difference between $\Delta_{\tau,1}$ and $\Delta_{\tau,2}$ therefore captures the potential bias introduced by comparing teachers observed in their second year rather than their first and is given by

$$\Delta_{\tau,2} - \Delta_{\tau,1} = \frac{\tau}{100}\left(\frac{\underline{a}-\overline{a}}{1-\underline{a}}\right)(q^B_H - q^A_H) \tag{3}$$

Equation (3) implies that the bias is zero if the attrition rate is constant across percentiles of college selectivity (i.e., if $\underline{a} = \overline{a}$) or if the upper boundary of the support of the distribution is unchanged by the licensing policy (i.e., if $q^B_H = q^A_H$).[35] If instead, $\overline{a} > \underline{a}$, then $\Delta_{\tau,2} > \Delta_{\tau,1}$ if and only if $q^B_H < q^A_H$, and $\Delta_{\tau,2} < \Delta_{\tau,1}$ if and only if $q^B_H > q^A_H$. In our context, we expect higher stringency to imply $q^B_H \leq q^A_H$ (see Proposition 1 of Larsen et al. 2020). In this case, $\Delta_{\tau,2} \geq \Delta_{\tau,1}$, meaning that using experienced teachers to measure the effect of higher stringency may overstate the true effect of the policy on the τth percentile among newly entering teachers.

Now consider the effect on a relatively high percentile τ of the distribution, such as the 90th percentile, such that $\min\{q^A_{\tau,2}, q^B_{\tau,2}\} \geq \phi$. For such a τ, the differences in the τth percentile for first-year and second-year teachers are given by:

$$\Delta_{\tau,1} \equiv q^B_{\tau,1} - q^A_{\tau,1} = (q^B_H - q^A_H) - \frac{100-\tau}{100}(q^B_H - q^A_H - q^B_L + q^A_L)$$

$$\Delta_{\tau,2} \equiv q^B_{\tau,2} - q^A_{\tau,2} = (q^B_H - q^A_H) - \frac{100-\tau}{100}\left(q^B_H - q^A_H - \frac{1-\underline{a}}{1-\overline{a}}(q^B_L - q^A_L)\right).$$

[35]More generally, the bias will be small if the attrition rate varies little across the distribution (i.e., if $\underline{a}$ is close to $\overline{a}$) or if the policy has little effect on the upper boundary of the support (i.e., if q^B_H is close to q^A_H).

The difference between $\Delta_{\tau,1}$ and $\Delta_{\tau,2}$ is therefore

$$\Delta_{\tau,2} - \Delta_{\tau,1} = \frac{100-\tau}{100}\left(\frac{\overline{a}-\underline{a}}{1-\overline{a}}\right)(q_L^B - q_L^A). \tag{4}$$

Equation (4) shows that the bias is zero if the attrition rate is constant across percentiles or if the left boundary of the support is unaffected by the policy change (i.e., if $q_L^B = q_L^A$). If instead, $\overline{a} > \underline{a}$, then $\Delta_{\tau,2} > \Delta_{\tau,1}$ if and only if $q_L^B > q_L^A$, and $\Delta_{\tau,2} < \Delta_{\tau,1}$ if and only if $q_L^B < q_L^A$. Proposition 1 of Larsen et al. (2020) predicts that higher stringency implies $q_L^B \geq q_L^A$. In that case, $\Delta_{\tau,2} \geq \Delta_{\tau,1}$, again implying a potential bias toward more positive estimates. Thus, for upper percentiles, estimation using experienced cohorts may understate the decline in the upper tail.

These arguments extend immediately to cohorts observed after more than two years of teaching. The longer a cohort has been teaching, the greater the potential bias toward positive estimated effects.

Our empirical analysis in the body of the paper includes some experienced-cohort comparisons such as those described above. These comparisons do not introduce bias if the attrition rate is relatively constant across the distribution. To examine this empirically, we modify equation (1) to include an interaction between $Stringency_{st}$ and the number of years a teacher cohort has been teaching when observed in a SASS survey. The results are reported in Appendix Table A.8. In panel A, stringency is interacted with a linear trend representing the number of years teaching for a given cohort, while in panel B, stringency is interacted with dummy variables for the number of years teaching. The columns correspond to those in Table 6: columns 1–3 report effects on the mean, columns 4–6 report effects on the 10th percentile, and columns 7–9 report effects on the 90th percentile. For each outcome, we present results separately for all teachers, elementary school teachers, and secondary school teachers. All regressions include the full set of state-by-year controls from our preferred specification (column 7 of Table 3).[36]

In both panels, we do not observe significant differences across cohorts in the effects of stringency on the distribution among secondary school teachers (columns 3, 6, and 9). We also find that the estimated coefficient on $Stringency_{st}$ for the left tail of the distribution

[36] Panel A does not include a main effect for the linear trend in years teaching, and panel B does not include main effects for the indicators of years teaching, because these are absorbed by year fixed effects. In panel B, the omitted years-teaching category corresponds to first-year teachers.

among secondary school teachers is similar to our main estimates in Table 3. These results suggest that the potential bias associated with teacher attrition described above does not drive our primary findings for secondary teachers. For elementary school teachers, however, column 2 of panels A and B indicates that the effect of stringency on the average is statistically larger for more experienced cohorts, and column 1 of panel B shows a similar pattern for the most experienced cohort in the all-teachers sample. This pattern is consistent with the attrition bias described above. We do not view this as a concern for our main conclusions, however, since our analysis in the body of the paper does not find robust effects for elementary school teachers.

A separate issue is that attrition rates as a function of college selectivity may differ across states. In that case, states could exhibit different trends in the τth percentile across cohorts even in the absence of policy changes. This would only bias our estimates if attrition-driven changes coincided systematically with changes in licensing stringency. The event study design in Section 4.1 addresses this concern by testing whether pre-trends differ across treated and untreated states. As an additional check, Appendix Table A.9 replicates column 8 of Table 3 but includes linear state-specific cohort trends (years teaching for a given cohort) rather than state-specific time trend across calendar years. The results are similar to those in Table 3, suggesting that cross-state variation in attrition rates does not drive our findings.

C Student Test Scores

We also provide some exploratory evidence on the effects of licensing stringency on the distribution of student test scores. We incorporate student-level test score data for the NAEP math and reading tests from NCES.[37] There is no way to link these test score data to the SASS data (for example, to construct value-added measures), but the test scores allow us to evaluate the effects of licensing stringency on student outcomes at the state-year level. The math test was administered in 1992, 1996, 2000, 2003, 2005, and 2007 to fourth and

[37] NAEP data are the only test scores available for a large set of states and years. The data do not contain information on students' baseline scores prior to being taught by a given teacher and do not contain a sufficient number of observations per teacher to permit the construction of standard measures of teachers' value-added. As such, NAEP data are not appropriate for constructing value-added measures, and for this reason, value-added analyses in the literature have been limited to case studies within single states. Such a case study would not suffice for our analysis, which exploits variation across all states and years in licensing stringency changes (as in Figure 3). We therefore use the NAEP data to analyze what might be called a *state* value-added model, asking whether a given state's distribution of student test scores improves after the state adopts stricter licensing rules.

eighth graders. The reading test was administered in 1992, 1994, 1998, 2002, 2003, 2005, and 2007 to fourth graders and, in those latter five years, to eighth graders. We normalize student test scores by subtracting the mean and dividing by the standard deviation across all students within a given grade and subject. As with our other variables, we then compute statistics of these test scores at the state-year level.[38]

Early NAEP waves — those before 2000 — do not record the precise year a student's teacher began teaching, which is necessary for identifying which licensing requirements the teacher faced. Instead, these earlier waves record only bins for teacher experience, including a bin for 0–2 years of experience. We address this issue by treating all students with teachers in the 0–2 years of experience bin in any pre-2000 NAEP wave as though they were taught by teachers with exactly one year of experience. For later waves, where we do not face this binning, we keep students taught by teachers with one year of experience. Descriptive statistics for our sample are reported in Appendix Table A.10. The underlying data includes math scores of 45,820 fourth graders and 44,810 eighth graders and reading scores of 51,070 fourth graders and 30,120 eighth graders. Because the NAEP test was administered only in certain years, we have fewer state-year cells available for this analysis.

We estimate regression (1) but replace the outcome variable with the average, 10th percentile, or 90th percentile of NAEP student test scores within a given state-year cell. We perform this exercise separately for fourth and eighth grade math and reading scores. The results are reported in Appendix Table A.11. In Appendix Table A.12, we repeat this exercise but with an alternative data contruction procedure, treating students in pre-2000 NAEP waves taught by teachers in the 0–2 years experience bin as though they are taught by teachers with two years of experience (rather than one), and, for post-2000 waves, keeping students taught by teachers with two years of experience. The qualitative implications of the two tables are similar.

In Appendix Table A.11, we find largely insignificant effects on the distribution of student

[38] As with the SASS data, we use individual weights to construct the overall mean and standard deviation and to construct statistics at the state-year level. NAEP does not report a raw test score but rather five plausible test score draws based on an Item Response Theory model; we use the average of these draws as a student's score. For any figures or tables using NAEP data, the full citation of the data is: U.S. Department of Education, National Center for Education Statistics, National Assessment for Education Progress (NAEP) Math Grades 4 and 8 National data for 2003, 2005, 2007 and State data for 1992, 1996, 2000; Reading Grades 4 and 8 National data for 2003, 2005, 2007 and State data for 1998 and 2002, as well as Reading Grade 8 State data for 1992 and 1994.

test scores.[39] We do observe statistically significant negative effects on the 90th percentile for grade 8 math score distribution (column 3), but these results are not significant at the 0.05 level in column 4, where state-year controls are included, nor in our alternative analysis in Appendix Table A.12. We therefore do not place much weight on this finding. In Appendix Table A.12, we observe a stastically significant negative effect on the 90th percentile for eight grade reading. Other point estimates in these tables, while not significant at conventional levels, are negative. If anything, the results are consistent with a decline in the student test score distribution (particularly in the right tail) under stricter licensing. However, we do not view these results are conclusive given that nearly all estimates are not measured precisely.

We emphasize that, given the limitations of the NAEP data, these results should be viewed as exploratory. A fruitful avenue for future research would be to further investigate the effects on student test scores as well as on non-cognitive outcomes, exploiting the various state-level licensing stringency changes identified in our study but focusing on case studies of individual states where richer test score data may be available.

D Cross-State Mobility

Teachers are required to satisfy the licensing requirements in the state in which they work, and it is possible that changes in licensing stringency may affect a teacher's choice of which state to work in. For example, college students planning to become teachers may choose a state in which to teach based on the ease of that state's licensing requirements. This type of cross-state mobility could potentially alter the college selectivity distribution in a treated state (one that increases stringency) as well as in a control state (one that leaves stringency unchanged). For example, suppose two states are initially equivalent — in stringency level, teacher pay, the college selectivity distribution, and any other aspect — and suppose one of these states (call it the *treated* state) increases its stringency while the other does not, and labor is mobile across states. Some potential teachers in the upper and lower tails may be driven out of teaching in the more stringent state and instead become teachers in the less stringent state, widening the distribution in the less stringent state and making the relative comparison of the college selectivity distribution between the treated and untreated states even starker. This suggests that part of the effects we measure in the body of the paper may

[39] The number of observations is smaller in columns that include all controls (even columns) because of the minimum cell-size requirement for SASS data, from which some controls are constructed.

be driven by such cross-state moves.

We analyze this issue empirically in two ways. First, we examine whether licensing stringency affects the fraction of teachers who report that they are currently working in a state different from the one in which they attended college. We construct this fraction in each state-year cell using SASS data combined with information on university locations from U.S. Department of Education (2019). On average, this fraction is 23%, suggesting that most teachers work in the same state in which they attended college. Treating this 23% level as the *natural level* of in-migration for teachers, we then examine whether this fraction changes systematically when a state increases its licensing stringency. In Appendix Table A.13.A, we present results from regressions analogous to those in equation (1), but with this fraction as the dependent variable. We find no significant effects of stringency on this outcome, and the confidence intervals are small enough to rule out economically meaningful effects.

Second, we examine the frequency of in-migration of teachers into a given state who were previously teaching in a different state. The SASS data suggest that such moves are rare to begin with: only 2% of teachers report having taught in a different state in the previous year. In Appendix Table A.13.B, we use this fraction as the dependent variable and the stringency in the SASS survey year as the right-hand-side variable of interest.[40] We find negative point estimates that are statistically significant, consistent with the possibility that in-migration declines when a state increases licensing restrictions.

In panel C, we examine whether states with higher in-migration rates differ in their college selectivity distribution. Here we estimate regressions similar to those in equation (1), but with the fraction of movers into the state (from panel B) as the right-hand-side variable of interest rather than stringency. Note that the estimates in panel C represent the effect of this fraction changing from 0 (no movers) to 1 (100% of teachers being movers). To interpret these estimates in percentage-point terms, they must be divided by 100. For example, the point estimate on the 10th percentile in column 3 suggests that a 1 percentage point increase in the fraction of movers (which corresponds to a 50% increase in the move rate relative to the mean) is associated with an increase of 0.005 standard deviations in the left tail for secondary school teachers. This magnitude is small and statistically insignificant. We also find no significant relationships in the other specifications in panel C. Thus, while a state's

[40] We use stringency in the survey year because a state's licensing requirements in that year are those that a newly arriving teacher would need to satisfy. Similarly, for any non-SASS controls in this regression (such as controls from the CPS), we use the values corresponding to the same SASS survey year.

stringency level appears to affect the rate of teacher in-migration, it does not appear to affect the distribution of college selectivity. Given these results, and given that the fraction of teachers who cross state borders after they begin teaching is small to begin with, we do not believe that cross-state migration in response to stringency poses a major concern for the interpretation of our main results.

This finding is consistent with several other papers showing that teachers rarely move across states. Strauss et al. (2000), studying Pennsylvania teachers, documented that 80% of teachers work within 70 miles of where they attended college, and 40% teach in the same district where they attended school as a child. Goldhaber et al. (2015), studying teachers in Washington and Oregon, showed that teachers are unlikely to move across states, and argued that this may be attributable to pension structures (i.e., teachers have already invested in a state's pension program) or to state licensing requirements that teachers have already invested in satisfying in their current state. Kim et al. (2017) reported similar findings. This latter argument suggests that, if anything, stricter licensing requirements in a given state would reduce cross-state mobility, consistent with findings in the broader occupational licensing literature (Johnson and Kleiner 2020). One study (Hirsch 2001) documented a case of teachers moving across state borders for higher pay.

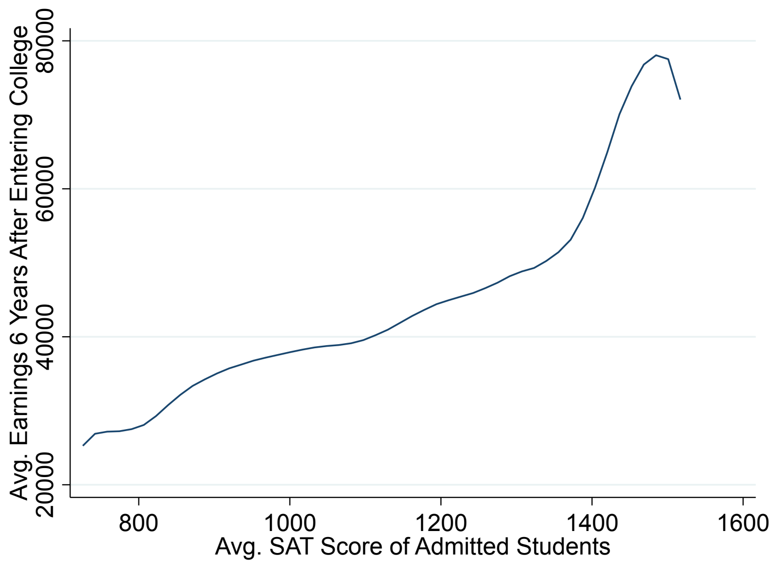
E Effects on Teacher Racial and Ethnic Diversity

We now investigate whether increased licensing stringency has any negative impact on the diversity of teacher supply, as measured by teacher race. We are motivated by previous findings in the education literature that certain forms of teacher certification requirements may negatively affect the entry of minority teachers, as well as by mixed findings in the broader occupational licensing literature showing both negative and positive effects of stricter licensing on minorities. Appendix Table A.14 shows the racial composition of teachers in our sample, with 82% of teachers being white, 7% Hispanic, 6% black, and 5% Asian or other non-white races.

Appendix Table A.15 presents the effects of licensing stringency on teachers' racial composition, estimated from a two-way fixed effects model similar to equation (1), but with the outcome variable replaced by the fraction of black, Asian, Hispanic, or non-white teachers in each state-year cell. We find a significant positive impact of increased stringency on the fraction of Hispanic elementary school teachers. The effect size, about 0.013–0.017, is economically meaningful given that the average fraction of Hispanic teachers is 7%.[41] Overall, we find no significant negative effects of increased stringency on minority participation.

F Additional Tables and Figures

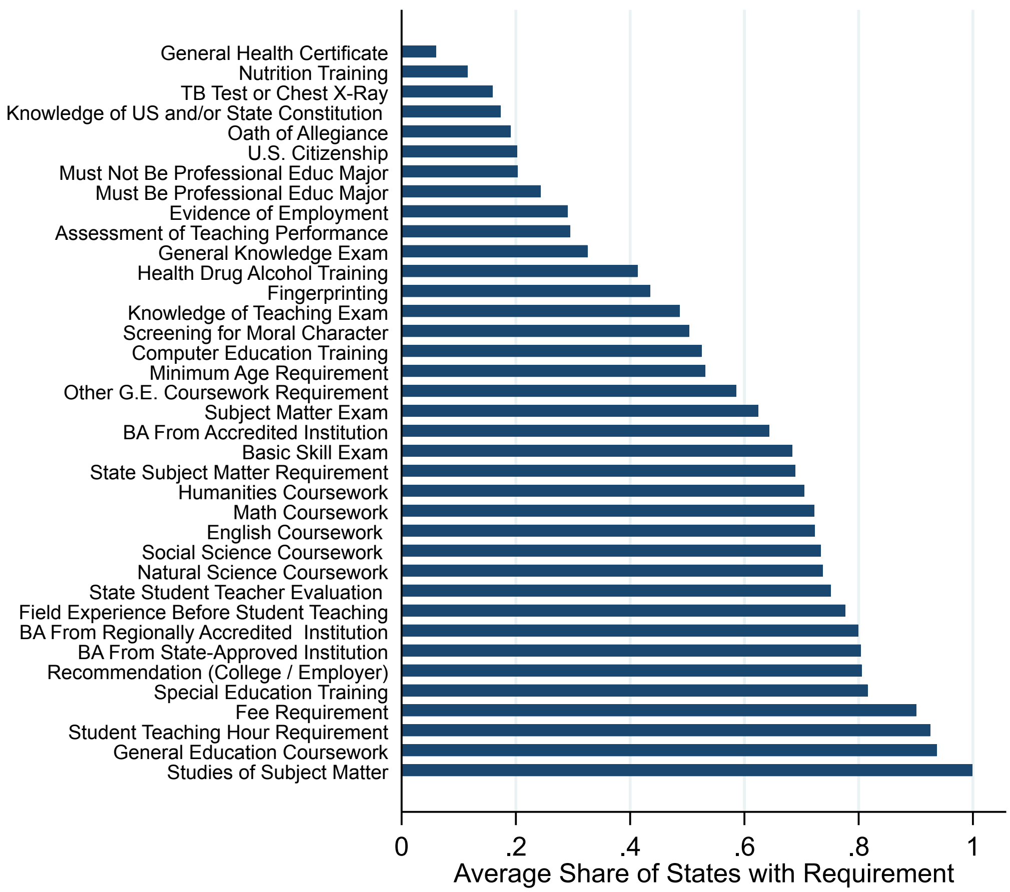
Figure A.1: Average SAT Score vs Earnings by Undergraduate Institution



Notes: This figure displays a local quadratic regression fit between average earnings (in 2007 dollars) six years after entering college (vertical axis) and the average SAT score of admitted students (horizontal axis). Each observation represents an undergraduate institution in the `collegescorecard.ed.gov` data.

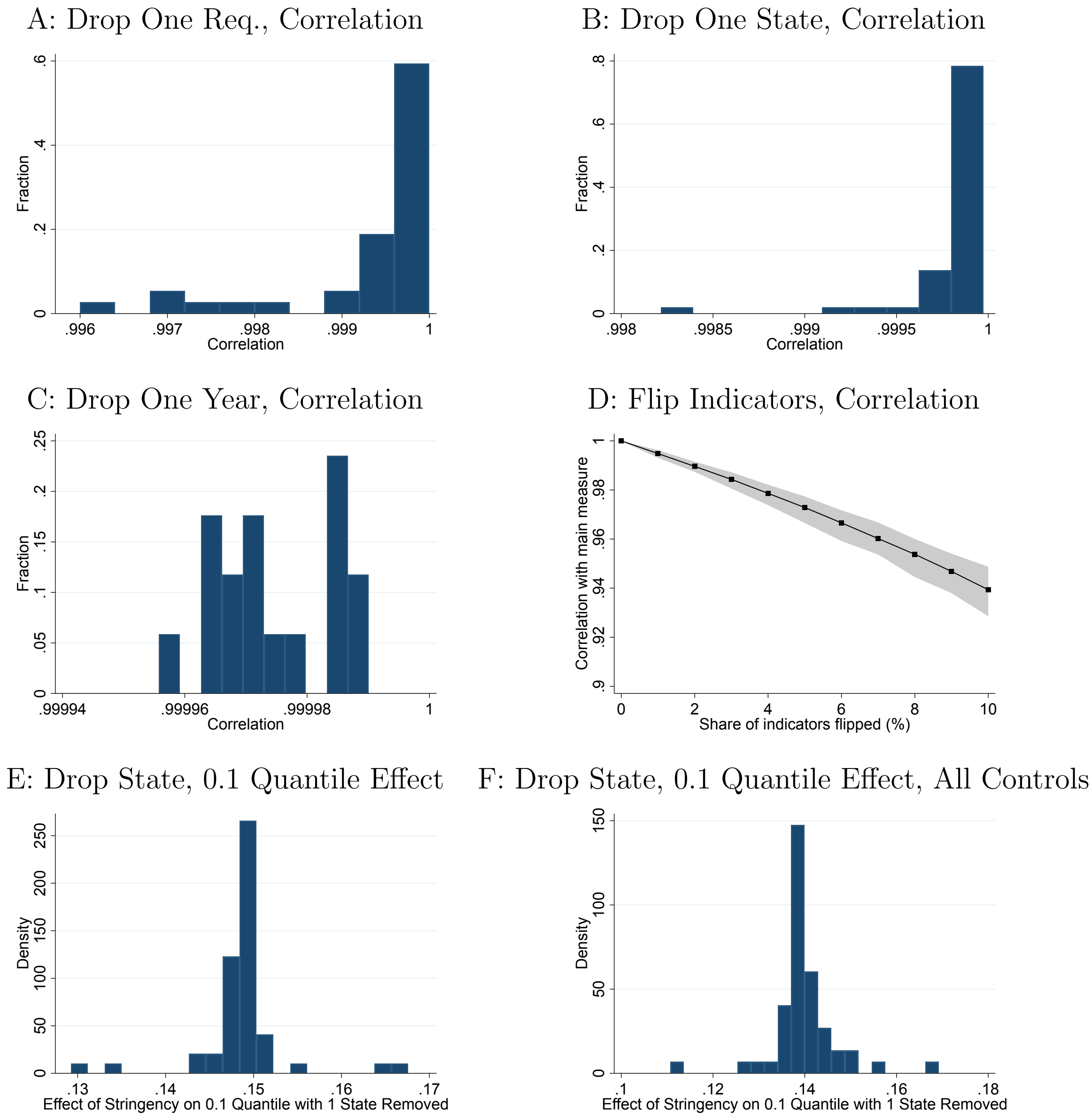
[41]This result differs from the negative effect on the fraction of minority participants documented in Angrist and Guryan (2008) or Deneault et al. (2026). Relative to these studies, we focus on different years and different aspects of licensing (coursework requirements rather than certification tests).

Figure A.2: Average Share of States with Each Requirement



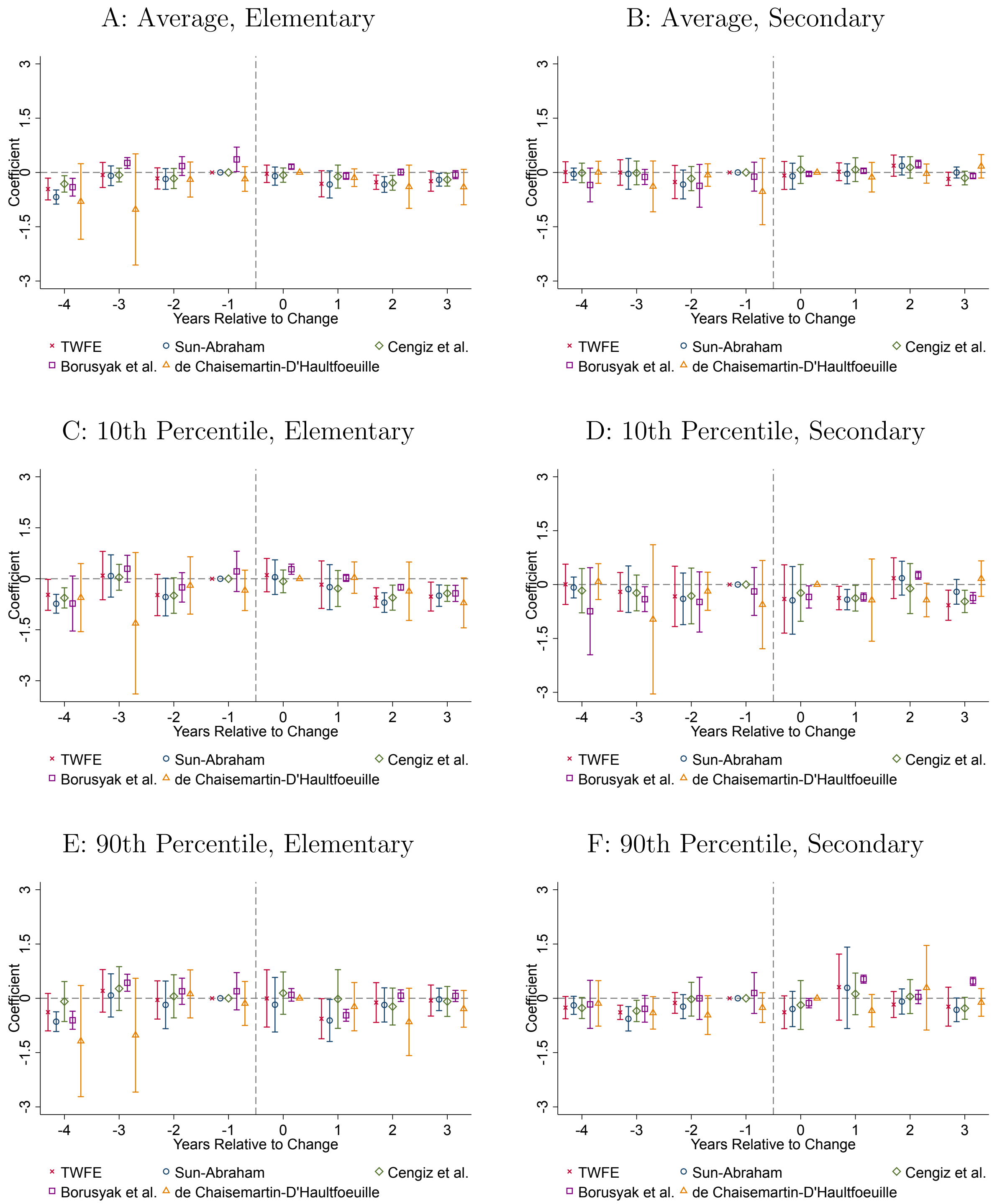
Notes: For each of the 37 licensing requirements, the figure shows the fraction of state-year pairs in which the requirement is in effect.

Figure A.3: Evaluating Potential Measurement Error in Stringency



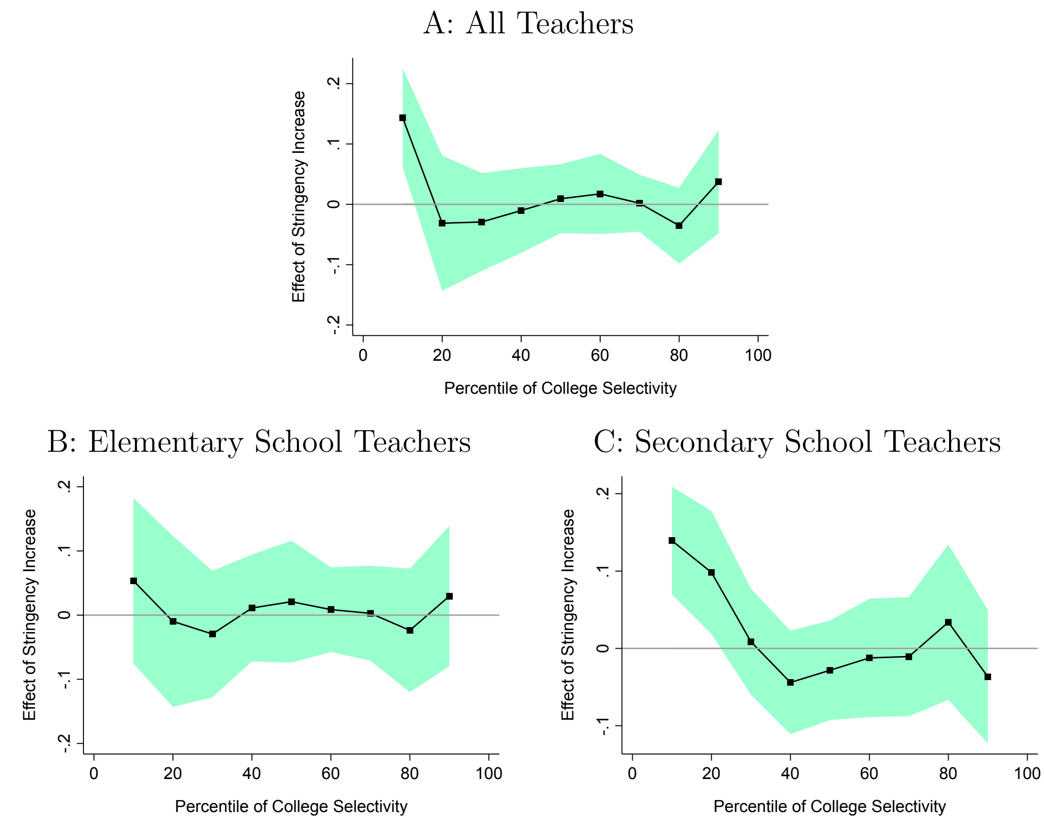
Notes: We evaluated whether measurement error in the 37 licensing requirements could affect the stringency measure. First, we recompute stringency after dropping each of the 37 requirements one at a time and calculate the correlation of the resulting measure with our main stringency measure. Panel A shows the histogram of these correlations. In panel B, we instead drop one state at a time, and in panel C we drop one year at a time. In panel D, we randomly change a fraction (ranging from 0.01 to 0.1) of state-year-requirement observations from 0 to 1 or vice versa, re-estimate stringency, and compute the correlation between the new measure and the main stringency measure. We repeat the randomization 50 times and plot the minimum, average, and maximum correlations. Panels E and F show histograms of the estimated effect of licensing stringency on the 10th percentile of college selectivity for secondary school teachers from equation (1), leaving out one state at a time. Panel E regressions include only state and year fixed effects, while panel F regressions include all controls from the preferred specification.

Figure A.4: Event Study for Elementary and Secondary School Teachers, High-to-Low States



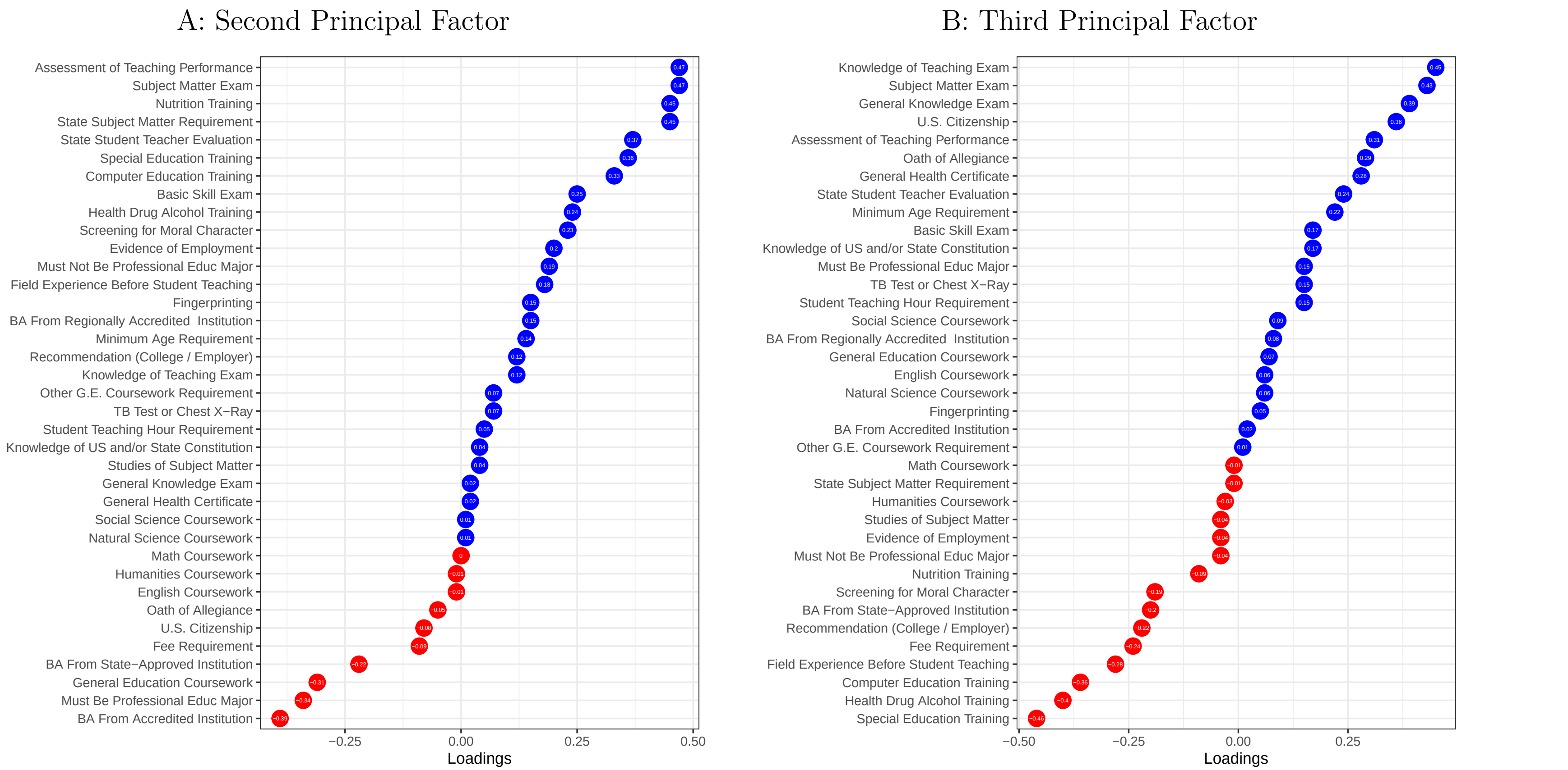
Notes: This figure presents event study estimates, using various methodologies, of the effect of a change from high to low licensing stringency on the average (panels A–B), 10th percentile (panels C–D), or 90th percentile (panels E–F) of college selectivity within a state-year. The left panels show results for elementary school teachers and the right panels show results for secondary school teachers. 95% confidence intervals are constructed with clustering at the state level.

Figure A.5: Effects of Licensing Stringency on the Full Distribution of College Selectivity



Notes: This figure presents the effect of increasing licensing stringency on the full distribution of teachers' college selectivity, estimated from equation (1). The outcomes are the 10th, 20th, ..., 90th percentiles of teachers' college selectivity. Panel A uses the full sample of teachers, panel B uses elementary school teachers, and panel C uses secondary school teachers. 95% confidence intervals for each percentile are shown by the shaded region and are constructed using seemingly unrelated regression to estimate all quantiles simultaneously with state-level clustering.

Figure A.6: Loadings of Certification Requirements on Second and Third Principal Factors



Notes: This figure displays the factor loadings of each of the 37 licensing requirements on the second principal factor (panel A) and the third principal factor (panel B).

Table A.1: Full Estimates from Table 3, Col 7, Panel A (All Teachers)

	(1) Average q	(2) 10th Percentile q	(3) 90th Percentile q
Stringency	0.0167	0.1433***	0.0374
	(0.0281)	(0.0445)	(0.0464)
Fraction Urban Schools	0.3904***	0.6988**	0.1622
	(0.1101)	(0.3082)	(0.1921)
Fraction Suburban Schools	0.2715***	0.3310	0.3453
	(0.1012)	(0.2468)	(0.2105)
Log Total Enrollment	-0.2317	-0.1690	-0.2733
	(0.2811)	(0.6689)	(0.6591)
Log State Educ. Expenditure	0.2345	0.1527	0.0036
	(0.2547)	(0.5435)	(0.4899)
Log No. Charter Schools	-0.0066	-0.0155	-0.0233
	(0.0111)	(0.0304)	(0.0205)
Average Percent Free Lunch	-0.0051**	-0.0042	-0.0025
	(0.0019)	(0.0041)	(0.0030)
Average Percent Minority Enrollment	-0.0009	-0.0073**	0.0035
	(0.0016)	(0.0036)	(0.0034)
Log Pub. Teacher Earnings (IPUMS)	-0.0870	-0.2281	0.0070
	(0.0986)	(0.2114)	(0.2002)
Log Pri. Teacher Earnings (IPUMS)	-0.0429	0.0348	-0.0968
	(0.0305)	(0.0550)	(0.0722)
Log Pri. Teacher Earnings (SASS)	-0.0017	-0.0102	0.1397
	(0.0740)	(0.1550)	(0.1466)
Pri. Teacher Earnings Exists (SASS)	0.0045	0.0964	-1.4378
	(0.7532)	(1.5818)	(1.4818)
Log Pub. Teacher Earnings (SASS)	0.3582*	0.4928	-0.0535
	(0.1991)	(0.4658)	(0.3687)
Log District BA Salary (SASS)	0.0429	-0.4567	-0.1073
	(0.5090)	(0.8947)	(1.0406)
Log District MA Salary (SASS)	-0.3066	-0.3450	0.4352
	(0.3584)	(0.8596)	(0.8459)
Fraction Teachers in Union (SASS)	-0.2322*	-0.4930*	-0.4253
	(0.1196)	(0.2884)	(0.3049)
Bartik Shock	-0.0617	-0.2815	0.6316
	(0.3664)	(0.7018)	(0.7580)
Log Per-capita Income	0.7509	1.8717	0.7011
	(0.5808)	(1.2811)	(1.0495)
Log Average Wage Income	-0.1500	-0.2038	-0.2706
	(0.1935)	(0.3263)	(0.4932)
Party of Governor	0.0806***	0.0399	0.0584
	(0.0190)	(0.0464)	(0.0532)
Party of State Legislature	0.0163	0.1534	-0.1327
	(0.0601)	(0.1189)	(0.0934)
Governor is Lame Duck	0.0013	0.1134*	-0.0488
	(0.0238)	(0.0566)	(0.0405)
Democrat-Republican Gov. Vote	-0.0020***	0.0006	-0.0032*
	(0.0007)	(0.0020)	(0.0016)
Post Financial Adequacy Policy	-0.0174	0.0792	0.0923
	(0.0400)	(0.0883)	(0.0806)
Log No. Desegregation Orders	-0.1225***	-0.1074	-0.1721**
	(0.0400)	(0.0714)	(0.0654)
Collective Bargaining Required	-0.2781	0.5536	-2.1868**
	(0.4958)	(1.0293)	(0.9002)
Collective Bargaining Allowed	-0.3503	0.1807	-2.1956**
	(0.4928)	(0.9906)	(0.8835)
Observations	857	857	857

Notes: This table reports most coefficient estimates from column 7 of panel A of Table 3. The full set of unemployment rate controls is omitted to save space. Each column represents a different regression, where the outcome is a different statistic of the distribution: the mean (column 1), the 10th percentile (column 2), or the 90th percentile (column 3). Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.2: Full Estimates from Table 3, Col 7, Panel B (Elementary School Teachers)

	(1) Average q	(2) 10th Percentile q	(3) 90th Percentile q
Stringency	0.0049	0.0534	0.0296
	(0.0404)	(0.0710)	(0.0602)
Fraction Urban Schools	0.3178*	0.4688**	0.4262
	(0.1791)	(0.2217)	(0.3385)
Fraction Suburban Schools	0.2724*	0.6351***	0.1167
	(0.1435)	(0.2306)	(0.2642)
Log Total Enrollment	0.4229	1.4088	1.3371*
	(0.4952)	(0.9607)	(0.7249)
Log State Educ. Expenditure	-0.1164	-0.3423	-0.3538
	(0.4614)	(0.7677)	(0.5789)
Log No. Charter Schools	-0.0059	-0.0133	0.0134
	(0.0202)	(0.0558)	(0.0322)
Average Percent Free Lunch	-0.0063**	-0.0002	-0.0153***
	(0.0026)	(0.0034)	(0.0045)
Average Percent Minority Enrollment	-0.0008	-0.0070**	0.0010
	(0.0020)	(0.0034)	(0.0031)
Log Pub. Teacher Earnings (IPUMS)	-0.0449	-0.5068	0.1875
	(0.1550)	(0.3481)	(0.3123)
Log Pri. Teacher Earnings (IPUMS)	-0.0684	-0.0930	-0.1873
	(0.0579)	(0.1042)	(0.1139)
Log Pri. Teacher Earnings (SASS)	-0.0166	0.0939	-0.0832
	(0.0962)	(0.1839)	(0.2017)
Pri. Teacher Earnings Exists (SASS)	0.1739	-1.0262	0.8041
	(0.9800)	(1.8759)	(2.0608)
Log Pub. Teacher Earnings (SASS)	0.0227	0.4850	0.0171
	(0.3326)	(0.3157)	(0.6556)
Log District BA Salary (SASS)	-1.0066*	-1.7311*	-2.4808*
	(0.5999)	(1.0091)	(1.3617)
Log District MA Salary (SASS)	0.9179	1.5619	2.3075
	(0.8171)	(1.2701)	(1.5253)
Fraction Teachers in Union (SASS)	-0.0787	-0.1064	-0.2303
	(0.0856)	(0.2222)	(0.1515)
Bartik Shock	0.0816	-0.3575	0.7180
	(0.5966)	(0.8286)	(1.1901)
Log Per-capita Income	1.0683	3.2071*	0.9420
	(0.9356)	(1.7386)	(1.6484)
Log Average Wage Income	-0.4654*	-0.8838*	-1.2738*
	(0.2586)	(0.4696)	(0.6548)
Party of Governor	0.1464***	0.0984	0.2039**
	(0.0374)	(0.0953)	(0.0809)
Party of State Legislature	0.0306	0.0712	0.1467
	(0.1064)	(0.1334)	(0.1699)
Governor is Lame Duck	-0.0044	-0.0010	-0.0190
	(0.0368)	(0.0629)	(0.0626)
Democrat-Republican Gov. Vote	-0.0037***	-0.0057*	-0.0068***
	(0.0010)	(0.0029)	(0.0017)
Post Financial Adequacy Policy	-0.0280	-0.0957	0.0091
	(0.0617)	(0.1521)	(0.0925)
Log No. Desegregation Orders	-0.1181**	-0.2272	-0.0864
	(0.0582)	(0.1734)	(0.1064)
Collective Bargaining Required	-0.1491	-1.8852	-1.2999
	(0.6548)	(1.7107)	(1.3291)
Collective Bargaining Allowed	-0.1440	-2.0242	-1.0350
	(0.6645)	(1.6581)	(1.3728)
Observations	696	696	696

Notes: This table reports most coefficient estimates from column 7 of panel B of Table 3. The full set of unemployment rate controls is omitted to save space. Each column represents a different regression, where the outcome is a different statistic of the distribution: the mean (column 1), the 10th percentile (column 2), or the 90th percentile (column 3). Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.3: Full Estimates from Table 3, Col 7, Panel C (Secondary School Teachers)

	(1) Average q	(2) 10th Percentile q	(3) 90th Percentile q
Stringency	0.0302	0.1394***	-0.0368
	(0.0271)	(0.0379)	(0.0469)
Fraction Urban Schools	0.2673*	0.6730***	-0.2819
	(0.1447)	(0.2403)	(0.2271)
Fraction Suburban Schools	0.2137*	0.5035**	-0.1571
	(0.1210)	(0.1993)	(0.1991)
Log Total Enrollment	-0.8185***	-1.2057**	-1.2797**
	(0.2795)	(0.4880)	(0.5306)
Log State Educ. Expenditure	0.4134**	0.6157*	0.3902
	(0.1862)	(0.3505)	(0.4283)
Log No. Charter Schools	-0.0037	0.0070	-0.0014
	(0.0146)	(0.0234)	(0.0215)
Average Percent Free Lunch	-0.0012	-0.0029	0.0060
	(0.0023)	(0.0059)	(0.0044)
Average Percent Minority Enrollment	-0.0037**	-0.0111*	0.0036
	(0.0015)	(0.0066)	(0.0031)
Log Pub. Teacher Earnings (IPUMS)	-0.0229	-0.1295	0.2545
	(0.0979)	(0.2239)	(0.2179)
Log Pri. Teacher Earnings (IPUMS)	-0.0353	-0.0085	-0.0361
	(0.0326)	(0.0658)	(0.0678)
Log Pri. Teacher Earnings (SASS)	0.0307	-0.1663	0.5013***
	(0.0758)	(0.1357)	(0.1252)
Pri. Teacher Earnings Exists (SASS)	-0.3373	1.7134	-5.1141***
	(0.7826)	(1.3800)	(1.2756)
Log Pub. Teacher Earnings (SASS)	-0.2363	-0.1023	-0.4038
	(0.2879)	(0.5502)	(0.6315)
Log District BA Salary (SASS)	0.4015	-0.9074	1.0574
	(0.5160)	(1.3656)	(0.8905)
Log District MA Salary (SASS)	0.1023	0.5955	-0.4748
	(0.4659)	(1.1557)	(0.9354)
Fraction Teachers in Union (SASS)	-0.1536	-0.2484*	-0.3780*
	(0.1247)	(0.1419)	(0.1992)
Bartik Shock	-0.6409	-0.9760	-0.9822
	(0.4566)	(0.9333)	(0.7609)
Log Per-capita Income	0.6449	0.8557	0.9760
	(0.5405)	(1.1066)	(1.0710)
Log Average Wage Income	0.2013	0.7363*	-0.0823
	(0.2365)	(0.4067)	(0.3828)
Party of Governor	0.0627**	0.0754*	0.0212
	(0.0267)	(0.0445)	(0.0668)
Party of State Legislature	-0.0139	0.1339	-0.0869
	(0.0459)	(0.0919)	(0.0901)
Governor is Lame Duck	0.0079	0.1366*	-0.0085
	(0.0275)	(0.0796)	(0.0473)
Democrat-Republican Gov. Vote	-0.0012*	0.0003	-0.0016
	(0.0006)	(0.0017)	(0.0017)
Post Financial Adequacy Policy	-0.0240	0.0990	-0.0387
	(0.0351)	(0.0720)	(0.0853)
Log No. Desegregation Orders	-0.1700***	-0.0711	-0.1793**
	(0.0412)	(0.0935)	(0.0830)
Collective Bargaining Required	-0.8880	-1.3287	-1.1334
	(0.6281)	(1.5367)	(1.2656)
Collective Bargaining Allowed	-0.8796	-1.6177	-1.0248
	(0.6422)	(1.5727)	(1.2711)
Observations	815	815	815

Notes: This table reports most coefficient estimates from column 7 of panel C of Table 3. The full set of unemployment rate controls is omitted to save space. Each column represents a different regression, where the outcome is a different statistic of the distribution: the mean (column 1), the 10th percentile (column 2), or the 90th percentile (column 3). Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.4: R^2 from Regressions in Table 3

A. All Teachers	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.827	0.836	0.830	0.830	0.828	0.831	0.847	0.860
10th Percentile: Stringency	0.682	0.689	0.686	0.688	0.683	0.688	0.708	0.730
90th Percentile: Stringency	0.550	0.554	0.556	0.562	0.551	0.555	0.580	0.607
B. Elementary	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.694	0.709	0.698	0.704	0.694	0.705	0.732	0.760
10th Percentile: Stringency	0.484	0.507	0.492	0.500	0.488	0.492	0.543	0.594
90th Percentile: Stringency	0.452	0.479	0.464	0.466	0.453	0.468	0.518	0.547
C. Secondary	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.802	0.810	0.804	0.805	0.804	0.803	0.820	0.835
10th Percentile: Stringency	0.652	0.668	0.656	0.658	0.654	0.660	0.687	0.706
90th Percentile: Stringency	0.538	0.544	0.554	0.550	0.539	0.538	0.575	0.599
State, Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics		Yes					Yes	Yes
Teacher Market Conditions			Yes				Yes	Yes
Non-teacher Market Conditions				Yes			Yes	Yes
Education Policy Controls					Yes		Yes	Yes
Political Conditions						Yes	Yes	Yes
State-Specific Trends								Yes

Notes: Each cell reports the R^2 from the regression underlying the corresponding cell in Table 3.

Table A.5: Effects of Licensing Stringency on College Selectivity Distribution, Balanced Panel

A. All Teachers	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.0179	0.0281	0.0163	0.0173	0.0193	0.0052	0.0167	0.0550**
	(0.0261)	(0.0311)	(0.0251)	(0.0261)	(0.0242)	(0.0291)	(0.0285)	(0.0241)
10th Percentile: Stringency	0.1546***	0.1512***	0.1467***	0.1292***	0.1525***	0.1652***	0.1463***	0.1828***
	(0.0447)	(0.0507)	(0.0431)	(0.0430)	(0.0432)	(0.0460)	(0.0449)	(0.0653)
90th Percentile: Stringency	0.0244	0.0331	0.0306	0.0522	0.0241	0.0093	0.0338	0.0338
	(0.0396)	(0.0477)	(0.0436)	(0.0359)	(0.0372)	(0.0457)	(0.0484)	(0.0539)
Observations	782	782	782	782	782	782	782	782
B. Elementary	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.0248	0.0972	0.0082	0.0267	0.0428	-0.0413	-0.0151	-0.0267
	(0.0850)	(0.0827)	(0.0947)	(0.1306)	(0.0672)	(0.0668)	(0.0543)	(0.0992)
10th Percentile: Stringency	0.1504	0.2018**	0.1335	0.1500	0.1674**	0.1617**	0.1539	0.1876
	(0.0842)	(0.0770)	(0.0808)	(0.1267)	(0.0679)	(0.0620)	(0.0910)	(0.1391)
90th Percentile: Stringency	-0.0852	0.0552	-0.1063	-0.1492	-0.0581	-0.2103*	-0.2829*	-0.3382
	(0.0793)	(0.0481)	(0.0868)	(0.1712)	(0.0849)	(0.0990)	(0.1480)	(0.2271)
Observations	153	153	153	153	153	153	153	153
C. Secondary	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	0.0257	0.0277	0.0321	0.0357	0.0281	0.0162	0.0335	0.0252
	(0.0266)	(0.0354)	(0.0307)	(0.0230)	(0.0259)	(0.0285)	(0.0306)	(0.0380)
10th Percentile: Stringency	0.1414***	0.1287***	0.1309***	0.1138**	0.1434***	0.1540***	0.1368***	0.1956***
	(0.0376)	(0.0466)	(0.0382)	(0.0452)	(0.0332)	(0.0402)	(0.0463)	(0.0594)
90th Percentile: Stringency	-0.0483	-0.0727	-0.0165	-0.0289	-0.0467	-0.0547	-0.0396	-0.0106
	(0.0748)	(0.0799)	(0.0745)	(0.0480)	(0.0738)	(0.0764)	(0.0518)	(0.0827)
Observations	680	680	680	680	680	680	680	680
State, Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics		Yes					Yes	Yes
Teacher Market Conditions			Yes				Yes	Yes
Non-teacher Market Conditions				Yes			Yes	Yes
Education Policy Controls					Yes		Yes	Yes
Political Conditions						Yes	Yes	Yes
State-specific Trends								Yes

Notes: This table reports a replication of Table 3, restricting the sample to states with non-missing data for all 17 years and thus producing a balanced panel of state-year observations. This restriction results in 46 states for the analysis of all teachers, 9 states for elementary school teachers, and 40 states for secondary school teachers.

Table A.6: Descriptive Statistics on Quantity of Teachers

	N	Mean	SD	Min	Max
No. Teachers (thousand, CCD)	867	55.74	59.16	4.39	321.93
Log No. Teachers (CCD)	867	10.46	0.99	8.39	12.68
No. Education Majors (thousand, IPEDS)	305	17.66	15.89	1.07	88.09
Log No. Education Majors (IPEDS)	305	9.37	0.96	6.98	11.39

Notes: This table presents summary statistics on measures of teacher quantity. The number of teachers in each state-year cell is obtained from the Common Core of Data (CCD), and the number of education majors is obtained from the Integrated Postsecondary Education Data System (IPEDS). IPEDS data are available only for even years from 1996–2008.

Table A.7: Data Sources for Raw Teacher Certification Requirements

#	Requirement description	1991	1993	1994-95	1996-97	1998-99	2000	2001	2002	2003	2004	2005	2006	2007	Mean	Std Dev
1	General Education Coursework	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	NA	NA	0.937	0.030
2	Studies of Subject Matter	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	DOE	NA	0.999	0.005
3	Special Education Training	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	NA	NA	0.815	0.105
4	Health Drug Alcohol Training	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	NA	NA	0.413	0.127
5	Computer Education Training	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	NA	NA	0.525	0.177
6	Nutrition Training	C1	NA	B1	B1	B1	B1	B1	B1	B1	B1	NA	NA	NA	0.115	0.037
7	Subject Matter Exam	C2	NA	B2	B2	B2	B2	B2	B2	B2	B2	NA	157	161	0.624	0.208
8	General Knowledge Exam	C2	NA	B2	B2	B2	B2	B2	B2	B2	B2	NA	157	161	0.325	0.154
9	Knowledge of Teaching Exam	C2	NA	B2	B2	B2	B2	B2	B2	B2	B2	NA	157	161	0.487	0.224
10	Assessment of Teaching Performance	C2	NA	B2	B2	B2	B2	B2	B2	B2	B2	NA	157	161	0.295	0.228
11	U.S. Citizenship	C3	N	B3	B3	B3	B3	B3	B3	B3	B3	I	NA	I	0.202	0.134
12	Oath of Allegiance	C3	M	B3	B3	B3	B3	B3	B3	B3	B3	H	NA	H	0.190	0.077
13	Evidence of Employment	C3	E	B3	B3	B3	B3	B3	B3	B3	B3	G	NA	G	0.291	0.381
14	Recommendation (College / Employer)	C3	H	B3	B3	B3	B3	B3	B3	B3	B3	C	NA	C	0.805	0.149
15	Minimum Age Requirement	C3	NA	B3	B3	B3	B3	B3	B3	B3	B3	NA	NA	NA	0.532	0.017
16	Fee Requirement	C3	B	B3	B3	B3	B3	B3	B3	B3	B3	P	NA	Q	0.901	0.048
17	General Health Certificate	C3	Q	B3	B3	B3	B3	B3	B3	B3	B3	NA	NA	NA	0.060	0.066
18	TB Test or Chest X-Ray	C3	L	B3	B3	B3	B3	B3	B3	B3	B3	K	NA	K	0.159	0.277
19	Knowledge of US and/or State Constitution	C3	NA	B3	B3	B3	B3	B3	B3	B3	B3	NA	NA	NA	0.173	0.075
20	Basic Skill Exam	C3	I	B3	B3	B3	B3	B3	B3	B3	B3	NA	157	161	0.684	0.304
21	Fingerprinting	C3	G	B3	B3	B3	B3	B3	B3	B3	B3	J	NA	J	0.435	0.328
22	Screening for Moral Character	C3	NA	B3	B3	B3	B3	B3	B3	B3	B3	NA	NA	NA	0.503	0.247
23	BA From Accredited Institution	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.644	0.109
24	BA From Regionally Accredited Institution	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.799	0.096
25	Must Be Professional Educ Major	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.243	0.082
26	Must Not Be Professional Educ Major	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.203	0.094
27	BA From State-Approved Institution	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.804	0.057
28	English Coursework	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.723	0.081
29	Humanities Coursework	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.705	0.081
30	Social Science Coursework	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.734	0.073
31	Natural Science Coursework	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.737	0.090
32	Other G.E. Coursework Requirement	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.586	0.065
33	Math Coursework	C4	NA	B4	B4	B4	B4	B4	B4	B4	B4	NA	NA	NA	0.722	0.102
34	Field Experience Before Student Teaching	C4	NA	B7	B7	B7	B7	B7	B7	B7	B7	NA	NA	B4	0.776	0.137
35	Student Teaching Hour Requirement	NA	NA	B8	B8	B8	B8	B8	B8	B8	B8	NA	DOE	NA	0.925	0.068
36	State Student Teacher Evaluation	C9	NA	B9	B9	B9	B9	B9	B9	B9	B9	NA	NA	NA	0.751	0.075
37	State Subject Matter Requirement	NA	NA	B10	B10	B10	B10	B10	B10	B10	B10	NA	NA	NA	0.689	0.092

Notes: This table shows the sources for the raw teacher certification requirements for all states for each year. Cells with three numbers or a number followed by a letter (e.g., "B4" or "157") indicate the table from a physical NASDTEC manual (from 1991, 1994–95, 1996–97, 1998–99, 2000, 2001, 2002, 2003, and 2004) or online cached versions of these manuals from 2006–2007 (NASDTEC Knowledgebase 2006–2007). Note that NASDTEC produced only one manual every two years from 1994–1999. Cells with a letter indicate the corresponding code from a physical Teacher Certification Requirements manual (Goddard 1993 or Boydston 2005–2007). Cells with "DOE" indicate data from the U.S. Department of Education Secretary's Sixth Annual Report on Teacher Quality (Duncan and Madzelan 2009). We also consulted AACTE (1990, 1991, 1993–1996), Boydston (1995–1999, 2001–2010), Goddard (1983–1993), and other reports from the U.S. Secretary of Education. Cells with "NA" (or years not shown, namely 1992) indicate that no information is available for that year. The second-to-last column reports the mean of the requirement dummy across all states and years, and the last column reports the average, across states, of the within-state standard deviation over time.

Table A.8: Effects of Licensing Stringency on College Selectivity Distribution, By Years of Teaching

	Average College Selectivity			10th Percentile			90th Percentile		
A: Linear	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stringency	-0.0106	-0.0499	0.0476*	0.1740***	-0.0213	0.1846**	-0.0230	0.0269	-0.0238
	(0.0281)	(0.0426)	(0.0281)	(0.0551)	(0.0816)	(0.0773)	(0.0609)	(0.0747)	(0.0578)
Str.*Years Teaching	0.0103	0.0204**	-0.0067	-0.0115	0.0278	-0.0174	0.0228	0.0010	-0.0050
	(0.0073)	(0.0088)	(0.0089)	(0.0191)	(0.0195)	(0.0233)	(0.0162)	(0.0205)	(0.0171)
B: Dummies	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stringency	-0.0046	-0.0306	0.0294	0.1612***	0.0129	0.1280	0.0083	0.0220	-0.0314
	(0.0314)	(0.0457)	(0.0278)	(0.0454)	(0.0981)	(0.0828)	(0.0602)	(0.0812)	(0.0535)
Str.*(Years Teach == 2)	0.0308	0.0246	0.0373	0.0309	0.0387	0.1133	0.0158	0.0176	-0.0098
	(0.0275)	(0.0343)	(0.0283)	(0.0470)	(0.0771)	(0.0971)	(0.0612)	(0.1059)	(0.0584)
Str.*(Years Teach == 3)	0.0073	0.0308	-0.0141	-0.0974	0.0005	-0.0324	0.0093	-0.0267	0.0283
	(0.0267)	(0.0297)	(0.0355)	(0.0600)	(0.0637)	(0.1000)	(0.0549)	(0.0722)	(0.0503)
Str.*(Years Teach == 4)	0.0424	0.0830*	-0.0193	-0.0184	0.1005	-0.0358	0.0727	0.0459	-0.0287
	(0.0383)	(0.0415)	(0.0378)	(0.0744)	(0.1048)	(0.1043)	(0.0759)	(0.0875)	(0.0853)
Str.*(Years Teach $\geq$ 5)	0.0412	0.0602	-0.0125	-0.0188	0.1064	-0.0060	0.1007	-0.0109	-0.0526
	(0.0280)	(0.0397)	(0.0380)	(0.0844)	(0.0748)	(0.0868)	(0.0679)	(0.1136)	(0.0708)
Observations	857	696	815	857	696	815	857	696	815
Sample	All	Elem	Sec	All	Elem	Sec	All	Elem	Sec
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents heterogeneous effects of licensing stringency on college selectivity, interacted with the number of years a teacher cohort has been teaching when observed in SASS data. The estimates are obtained by modifying equation (1) to include an interaction between stringency and the number of years a cohort has been teaching: a linear interaction in panel A and interactions with dummy variables for years of teaching in panel B. "Str." stands for stringency. Each column in each panel corresponds to a separate regression. The outcome is the mean (columns 1–3), the 10th percentile (columns 4–6), or the 90th percentile (columns 7–9) of college selectivity within the state-year cell. Columns 1, 4, and 7 report results for all teachers; columns 2, 5, and 8 restrict the sample to elementary school teachers; and columns 3, 6, and 9 restrict the sample to secondary school teachers. All regressions include the full set of state-by-year controls, as in column 7 of Table 3. Standard errors are clustered at the state level and reported in parentheses. $^*: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.9: Effects of Licensing Stringency on College Selectivity Distribution, Controlling for State-Specific Cohort Trends

	(1)	(2)	(3)
Average: Stringency	0.0062	0.0089	0.0090
	(0.0338)	(0.0444)	(0.0348)
10th Percentile: Stringency	0.1360^{**}	0.0866	0.1170^{***}
	(0.0528)	(0.0798)	(0.0420)
90th Percentile: Stringency	-0.0001	0.0546	-0.0973^{*}
	(0.0527)	(0.0718)	(0.0542)
Observations	857	696	815
Sample	All	Elem	Sec
All State-by-Year Controls	Yes	Yes	Yes
State-Specific Cohort Trends	Yes	Yes	Yes

Notes: This table presents results analogous to column 8 of Table 3, but with state-specific time trends replaced by state-specific cohort trends. Each cell in the table corresponds to a different regression, and the coefficient on *Stringency* is reported. Column 1 shows results for all teachers, column 2 restricts the sample to elementary school teachers, and column 3 restricts the sample to secondary school teachers. Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.10: Descriptive Statistics for NAEP Student Test Scores

	Grade 4 Math		Grade 8 Math		Grade 4 Reading		Grade 8 Reading	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Average Score	-0.02	0.29	-0.00	0.30	-0.01	0.27	0.00	0.25
10th Percentile Score	-1.28	0.38	-1.23	0.36	-1.30	0.37	-1.30	0.37
90th Percentile Score	1.21	0.25	1.21	0.31	1.19	0.23	1.16	0.25
Num. Obs. Per State-Year	163	99	160	94	158	93	128	76
Total Student Obs.	45,820		44,810		51,070		30,120	
Total State-Year Obs.	281		280		323		236	

Notes: This table shows summary statistics of NAEP student test scores.

Table A.11: Effects on Student Test Score Distribution

	Grade 4 Math		Grade 8 Math		Grade 4 Reading		Grade 8 Reading	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	-0.0243	-0.0476	-0.0284	-0.0093	-0.0154	-0.0347	-0.0437	-0.0447
	(0.0416)	(0.0449)	(0.0199)	(0.0253)	(0.0393)	(0.0271)	(0.0417)	(0.0320)
10th Percentile: Stringency	-0.0670	-0.0602	0.0149	0.0365	-0.1066*	-0.0870	-0.1146	-0.1171
	(0.0579)	(0.0648)	(0.0294)	(0.0408)	(0.0622)	(0.0583)	(0.0776)	(0.0843)
90th Percentile: Stringency	-0.0015	-0.0572	-0.0665***	-0.0485	0.0629*	0.0413	-0.0179	-0.0270
	(0.0209)	(0.0373)	(0.0198)	(0.0322)	(0.0355)	(0.0323)	(0.0151)	(0.0493)
State, Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls		Yes		Yes		Yes		Yes
Observations	281	229	280	265	323	268	236	228

Notes: This table presents the effects of licensing stringency on the distribution of student test scores, estimated from a two-way fixed effects model similar to equation (1). Each observation is a state-year cell. Each cell in the table corresponds to a different regression, and each row represents a different outcome. The coefficient on *Stringency* is reported. Columns 1–2 show results for grade 4 math scores, columns 3–4 for grade 8 math scores, columns 5–6 for grade 4 reading scores, and columns 7–8 for grade 8 reading scores. Odd columns include no additional controls, and even columns include all controls from the preferred specification, as in column 7 of Table 3. Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.12: Effects on Student Test Score Distribution, Alternative Data Construction

	Grade 4 Math		Grade 8 Math		Grade 4 Reading		Grade 8 Reading	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average: Stringency	-0.0060	0.0062	0.0124	0.0219	-0.0249	-0.0394	-0.0500	-0.0762*
	(0.0385)	(0.0429)	(0.0278)	(0.0298)	(0.0358)	(0.0305)	(0.0449)	(0.0430)
10th Percentile: Stringency	-0.0339	-0.0338	-0.0005	0.0271	-0.0686	-0.0756	-0.0793	-0.1481**
	(0.0492)	(0.0616)	(0.0681)	(0.0467)	(0.0626)	(0.0523)	(0.0874)	(0.0638)
90th Percentile: Stringency	-0.0112	-0.0038	0.0190	0.0589	0.0039	-0.0337	-0.0769*	-0.0697**
	(0.0333)	(0.0509)	(0.0288)	(0.0389)	(0.0236)	(0.0277)	(0.0429)	(0.0337)
State, Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All Controls		Yes		Yes		Yes		Yes
Observations	239	186	238	225	281	224	236	224

Notes: This table is similar to Table A.11 but based on a different sample and assumptions. This sample uses students from post-2000 NAEP waves who are taught by teachers with two years of experience. For students in pre-2000 waves, this sample treats students taught by teachers with 0–2 years of experience in NAEP as though they were taught by teachers with two years of experience. Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.13: Effect of Stringency on Teachers Moving States

A. Teach in Diff State than College	(1)	(2)	(3)
Stringency	-0.0079	-0.0234	0.0174
	(0.0078)	(0.0166)	(0.0161)
B. Teach in Diff State Last Year	(1)	(2)	(3)
Stringency	-0.0088***	-0.0096**	-0.0069**
	(0.0027)	(0.0039)	(0.0031)
C. College Selectivity Regressed on Fraction Movers	(1)	(2)	(3)
Average: Fraction Movers	0.0379	-0.3054	0.0591
	(0.4172)	(0.4185)	(0.4052)
10th Percentile: Fraction Movers	0.5034	-0.3633	0.5378
	(0.7828)	(0.6639)	(0.7291)
90th Percentile: Fraction Movers	-0.3337	0.2304	0.2940
	(0.5949)	(0.7664)	(0.5654)
Observations	857	696	815
Sample	All	Elem	Sec
All Controls	Yes	Yes	Yes

Notes: Panels A and B present the results of equation (1), where the left-hand-side variable is the fraction of teachers working in a state different from the one in which they attended their undergraduate institution in panel A and the fraction of teachers teaching in a different state than the one in which they taught in the previous year in panel B. Panel B uses stringency in the survey year as the right-hand-side variable of interest. Panel C shows regressions with the same outcomes as in Table 3, but with the right-hand-side variable of interest being the fraction of teachers in a state-year cell who taught in a different state in the previous year. Each cell in the table corresponds to a different regression. Column 1 reports results for all teachers, column 2 for elementary school teachers, and column 3 for secondary school teachers. All regressions include all state-year controls, as in the preferred specification (column 7 of Table 3). Standard errors are clustered at the state level and reported in parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Table A.14: Summary Statistics of Racial Composition of Teachers

	All Teachers				Elementary		Secondary	
	Mean	SD	Min	Max	Mean	SD	Mean	SD
Fraction Asian	0.02	0.06	0.00	0.90	0.02	0.06	0.02	0.06
Fraction Black	0.09	0.10	0.00	1.00	0.08	0.12	0.10	0.12
Fraction Hispanic	0.07	0.09	0.00	0.47	0.08	0.12	0.07	0.09
Fraction White	0.88	0.11	0.00	1.00	0.89	0.13	0.87	0.13
Total State-Year Observations	857				696		815	

Notes: This table presents summary statistics on teacher race.

Table A.15: Effects of Licensing Stringency on Racial Composition of Teachers

	(1)	(2)	(3)
Black: Stringency	-0.0095	-0.0139	-0.0072
	(0.0139)	(0.0254)	(0.0103)
Asian: Stringency	0.0009	-0.0004	0.0010
	(0.0035)	(0.0040)	(0.0031)
Hispanic: Stringency	0.0157***	0.0136*	0.0176**
	(0.0048)	(0.0077)	(0.0078)
Nonwhite: Stringency	-0.0109	-0.0185	-0.0068
	(0.0116)	(0.0227)	(0.0099)
Observations	857	696	815
Sample	All	Elem	Sec
All Controls	Yes	Yes	Yes

Notes: This table presents the effects of licensing stringency on teachers' racial composition, estimated from a two-way fixed effects model similar to equation (1). Each observation is a state-year cell. Each cell in the table corresponds to a different regression, and each row represents a different outcome: the fraction of Black, Asian, Hispanic, and non-white teachers in each state-year cell. The coefficient on *Stringency* is reported. Column 1 shows results for all teachers, column 2 restricts the sample to elementary school teachers, and column 3 restricts the sample to secondary school teachers. All regressions include all controls from the preferred specification, as in column 7 of Table 3. Standard errors are clustered at the state level and reported in the parentheses. $^{*}: p < 0.10$, $^{**}: p < 0.05$, and $^{***}: p < 0.01$.

Appendix References

AACTE (1990, 1991, 1993-1996). *Teacher Education Policy in the States: A 50-State Survey of Legislative and Administrative Actions*. American Association of Colleges for Teacher Education, Washington, DC. Six certification manuals.

Astin, A., Green, K. C., Korn, W. S., and Maier, M. J. (1983). *The American Freshman: National Norms for Fall 1983*. Cooperative Institutional Research Program of the American Council on Eduaction, Higher Education Research Institute, UCLA.

Boydston, J. E. (1995-2010). *Teacher Certification Requirements in All 50 States: How and Where to Get a Teaching Certificate*. Teacher Certification Publications, Sebring, FL. Fifteen cerfication manuals from 1995 to 1999 and from 2001 to 2010.

Duncan, A. and Madzelan, D. T. (2009). The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom. *U.S. Department of Education, Office of Postsecondary Education*.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *ECMA*, 77(3):953–973.

Goddard, R. E. (1983-1993). *Teacher Certification Requirements in All 50 States: How and Where to Get a Teaching Certificate*. Teacher Certification Publications, Grover Hill, OH. Eleven cerfication manuals from 1983 to 1993.

Goldhaber, D., Grout, C., Holden, K. L., and Brown, N. (2015). Crossing the Border? Exploring the Cross-State Mobility of the Teacher Workforce. *Educational Researcher*, 44(8):421–431.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *JHR*, 39(2):326–354.

Hirsch, E. (2001). Teacher Recruitment: Staffing Classrooms with Quality Teachers. ERIC Paper Number ED453199.

Johnson, J. E. and Kleiner, M. M. (2020). Is Occupational Licensing a Barrier to Interstate Migration? *American Economic Journal: Economic Policy*, 12(3):347–373.

Kim, D., Koedel, C., Ni, S., and Podgursky, M. (2017). Labor Market Frictions and Production Efficiency in Public Schools. *Economics of Education Review*, 60:54–67.

Kirby, S. N., Naftel, S., and Berends, M. (1999). Staffing At-Risk School Districts in Texas. Problems and Prospects. ERIC No. ED436496.

NASDTEC (1991, 1994, 1996, 1998, 2000–2004). *National Association of State Directors of Teacher Education Certification (NASDTEC) Manual on the Preparation and Certification of Education Personnel*. Kendall/Hunt Publishing Company, Dubuque, IA. Seven certification manuals.

NASDTEC Knowledgebase (2003–2010). *National Association of State Directors of Teacher Education Certification (NASDTEC) Manual on the Preparation and Certification of Education Personnel and NASDTEC Knowledgebase*. Kendall/Hunt Publishing Company, Dubuque, IA. As published on National Center for Education Statistics (NCES) website: http://nces.ed.gov/programs/digest/. Eight tables from years 2003-2010.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2018). IPUMS USA: Version 8.0 [Dataset]. Minneapolis: University of Minnesota. https://doi.org/10.18128/D010.V8.0.

U.S. Department of Education (2019). National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), IPEDS Data Tool. https://nces.ed.gov/ipeds/.