

# TEACHER ACCOUNTABILITY REFORMS AND THE SUPPLY OF NEW TEACHERS

Matthew A. Kraft  
*Brown University*

Eric J. Brunner  
*University of Connecticut*

Shaun M. Dougherty  
*University of Connecticut*

David Schwegman  
*Syracuse University*

January 2018

## Abstract

In recent years, states across the country have attempted to increase the accountability of public school teachers by implementing rigorous, high-stakes evaluation systems and in some cases repealing teacher tenure protections. We examine the effect of these reforms on the supply of new entrants into the teacher labor market by exploiting a unique panel dataset that includes the number of teaching licenses granted by states. Leveraging variation in the adoption of reforms across states and time, we find that evaluation reforms resulted in a steady decline in the statewide supply of new teachers, whereas tenure reforms produced a sharp but more temporary contraction. In exploratory analyses, we find no evidence that decreases in labor supply differed systematically across non-shortage and shortage licensure areas. We find mixed evidence of the effect of accountability on the selectivity of the institutions where prospective teachers earned their teaching degrees. Tenure reforms appear to have reduced supply most among candidates from less selective universities, while there is little evidence evaluation reforms had any differential effect by university selectivity.

JEL No. I21, I28, J22, J24

Correspondence regarding the paper can be sent to Matthew Kraft at [mkraft@brown.edu](mailto:mkraft@brown.edu); Brown University, PO Box 1938, Providence RI, 02478; (401) 863-3597. This work was generously supported by the Spencer Foundation [Award #201700052] and the Institute for Education Sciences [Award # R305A170053]. We thank Motaz Al-Chanati, Alif Al-Ibrahim, Michael Briskin, Bruna Lee, Matt Owen, and Sarah Novicoff for their excellent research assistance. We are grateful to David Deming, Tom Dee, Joshua Hyman, Joshua Goodman, Steve Ross, Matthew Steinberg, Eric Taylor and seminar participants at the University of Arkansas, University of Connecticut, and the Harvard Kennedy School who provided helpful comments.

## I. Introduction

In response to mounting public pressure and strong incentives from the federal government, state legislatures across the country have enacted laws aimed at increasing accountability for public school teachers. Many of these efforts have focused on implementing high-stakes teacher evaluation systems and repealing teacher tenure. By 2016, 44 states had implemented major reforms to their teacher evaluation systems. These reforms were intended to increase the rigor of the evaluation process and streamline the dismissal process. Six states enacted laws that effectively eliminated tenure for new teachers, with three additional state legislatures passing laws that were vetoed or immediately overturned by ballot referendum.

Proponents assert these reforms are necessary given the wide variation in teacher quality documented in the literature and the inability or reluctance of school districts to terminate ineffective teachers (Wiesberg et al., 2009; Hanushek, 2009). In contrast, opponents of accountability reforms argue that teachers are already under immense pressure to perform and that the implementation of high-stakes teacher evaluation systems and the elimination of tenure protections would only serve to reduce both the quantity and quality of new college graduates willing to pursue a career in teaching (Fullan, 2011). However, limited empirical evidence exists on whether and how those reforms actually affect the teacher labor market.

What evidence is available comes primarily from studies that examine the effect of changes in job security on the behavior and career decisions of *current* teachers. Research demonstrates that teachers are less likely to be absent, spend more of their own money on classroom materials, and participate in more extracurricular activities in the years before they earn tenure (Hansen, 2009, Jones, 2015).<sup>1</sup> Studies have also shown that the increased dismissal

---

<sup>1</sup> Similarly, Jacob (2013) finds that a policy change that gave Chicago Public School principals more freedom to dismiss probationary teachers reduced teacher absences by 10 percent.

threat under new high-stakes evaluation has led to greater voluntary attrition of lower-performing teachers (Dee and Wyckoff, 2015; Loeb, Miller and Wyckoff, 2014; Sartain and Steinberg, 2016; Cullen, Koedel, and Parsons, 2016; Strunk, Barret and Lincove, 2017).

This paper addresses the largely unexamined question of how accountability reforms affect the supply of *new* teachers. We exploit arguably exogenous variation in the timing of teacher accountability reforms across states to provide the first empirical evidence on how these reforms affected new teacher labor supply. Our analyses also provide a direct empirical test of a key assumption of the teacher quality literature, namely that accountability reforms do not affect the willingness of prospective teachers to enter the teacher labor market. Many prior studies estimate potential learning gains from dismissing low-performing teachers through simulation analyses that rely on the untested assumption that dismissed teachers can always be replaced with average-quality novice teachers (Gordon, Kane, and Staiger, 2006; Hanushek, 2009; Staiger and Rockoff, 2010; Winters and Cowen, 2013a, 2013b; Goldhaber and Hansen, 2010).

Our work builds most directly on Rothstein’s (2015) simulation analyses of how a policy that removes the bottom 20 percent of performers among teachers eligible for tenure would affect the supply of both current and future teachers. Although Rothstein’s policy simulation is far more extreme than any observed in practice, his findings suggest that it would require “substantial increases in teacher salaries” to offset the accountability policy and continue to draw equivalent numbers of new entrants into the teaching profession (p.126).

We begin by presenting trends in the supply of new teachers from 2002 to 2016 based on the total number of teaching licenses granted by states each year, information that states are required to report to the U.S. Department of Education (DOE). This outcome provides a direct measure of the potential supply of all new teachers eligible to work in publicly-funded schools

(traditional or charter), regardless of certification pathway or licensure type, and links each teacher to their state of intended employment. Obtaining a license is often the last step prior to entering the labor market, and thus is likely to be immediately responsive to policy reforms. We complement our licensure measure with data on the number of university-based teacher preparation program degree completers from the Integrated Postsecondary Education Data System (IPEDS). Both measures suggest that, nationally, new teacher labor supply has fallen by over 20 percent between pre-recession levels in 2007 and 2016.

We then provide evidence on the relationship between accountability reforms and the supply of new teachers using a completely flexible event-study specification that allows for an examination of pre- and post-treatment trends. We complement this non-parametric approach with more structured difference-in-differences (DD) models to increase the precision of our estimates. In our preferred models, we allow for differential linear trends pre- and post-policy reform in treated states to explicitly test the parallel trends assumption as well as to model separately any immediate effects of the policy reforms and any incremental effects on teacher labor supply over time.

We find consistent evidence that both implementing high-stakes evaluation reforms and repealing tenure reduced teacher labor supply. In joint models, we estimate that the number of licenses granted dropped by approximately 15 percent in states that adopted evaluation reforms and 16 percent in states that repealed tenure, on average. Flexible models suggest that the timing dynamics of these effects differ. Specifically, evaluation reforms appear to result in a steady decline in new labor supply over time whereas tenure reforms result in a sharp and immediate contraction in the supply of new teachers that then slowly rebounds. These findings are robust to

a range of model specifications and falsification tests and are consistent with Rothstein's (2015) predictions given that we find no evidence of offsetting increases in teacher salaries.

The consequences of these reductions in new teacher labor supply depend in large part on the composition of the prospective teachers who pursue other occupations. Reductions in supply could reflect market corrections to the large oversupply of general elementary and secondary humanities teachers. Reductions in supply could also reflect a disproportionate decrease in the number of less-qualified (or highly-qualified) prospective teachers. We explore these possibilities using more detailed IPEDS data on the number of graduates from university-based teacher preparation programs. These data provide a less precise measure of state-specific new teacher labor supply in the public sector, but allow us to examine whether the impact of accountability reforms varies by subject area specialization, the selectivity and average freshman SAT scores of institutions where prospective teachers attended a preparation program (a proxy for teacher quality), gender and race.

We find no evidence that decreases in labor supply differed systematically across non-shortage versus shortage licensure areas. Reductions were proportionally largest for secondary English (non-shortage) and secondary science (shortage). We find mixed evidence of differential effects of accountability reforms on the qualifications of graduates from teacher preparation programs based on institutional selectivity and average freshman SAT scores. Effects of high-stakes evaluation reforms on teacher qualifications are mixed and imprecisely estimated, while evidence suggests tenure reforms increased the quality of the institutions where prospective teachers receive their degrees. We also find no conclusive evidence that accountability reforms had a differential effect on prospective teachers by gender or race. We conclude with a discussion of the implications for policy, practice and future research.

## **II. Teacher Accountability Reforms**

Efforts to introduce greater accountability in schools and classrooms have ebbed and flowed throughout the history of U.S. public education. Present George W. Bush's No Child Left Behind (NCLB) act, passed in 2002, expanded test-based school accountability policies nationally and established more rigorous teacher licensure standards with the aim of improving teacher quality. In following years, a growing body of research on teacher effectiveness exploited new district administrative datasets linking students to teachers that NCLB helped to create. Three seminal findings from this research served as signposts for the teacher accountability reform initiatives of the Obama administration: 1) the effects teachers have on student learning are large and vary considerably across teachers, 2) teacher qualifications are weakly related to their effects on student learning, and 3) teacher evaluation systems were failing to differentiate among teachers despite their large differences in effectiveness.

Starting in 2009, the Obama administration leveraged \$4.35 billion from the American Reinvestment and Recovery Act to fund the Race to the Top (RTT) grant competition. The RTT application rubric detailed specific evaluation system reforms required for a competitive proposal such as evaluating teachers using multiple measures including student achievement growth, rating teachers on a scale with multiple categories, conducting annual evaluations and providing feedback, and using evaluations to make high-stakes personnel decisions. The opportunity to secure hundreds of millions of dollars in federal grant aid was too tempting for most states to pass up in the wake of the Great Recession. Forty-six states applied in the first of three rounds of funding under RTT. Private foundations and philanthropic organizations such as the Gates and Broad Foundations and the New Schools Venture Fund also invested millions of

dollars to support evaluation reforms across the country. In 2012, the Obama administration then made adopting evaluation reforms a condition for states to receive a waiver from the increasingly stringent accountability consequences of failing to meet targets set by NCLB.

By 2016, 44 states had passed legislation that mandated major teacher evaluation reforms (NCTQ, 2016). While new evaluation systems differ across states, nearly all systems share a common set of features: 1) the incorporation of multiple measures of teacher performance including test-based performance measures such as value-added measures or student growth percentiles; 2) the use of multiple performance rating categories; and 3) the use of evaluation ratings to inform high-stakes personnel decisions. In many cases, teachers receiving low evaluation ratings could be dismissed or denied tenure or promotion, while teachers with exemplary ratings could be rewarded with merit pay or promoted to new positions on a career ladder (Donaldson and Papay, 2015).

Over this same period, several states attempted to roll back teacher tenure protections, arguing that tenure laws made it unreasonably difficult to dismiss poor performing teachers. Teachers' organizations and unions had long worked to secure and protect state statutes that require school districts to follow detailed due process procedures and timelines when attempting to dismiss a tenured teacher. These statutes, in combination with locally bargained employment contracts and due process procedures, establish the binding set of job protections known as teacher tenure.<sup>2</sup> In practice, teachers typically earn tenure by default after two to five years of work experience. As of 2016, lawmakers in nine states had passed legislation aimed at

---

<sup>2</sup> By the turn of the 21st century, all but two states (OR, MS) had adopted formal tenure statutes.

eliminating or effectively removing formal teacher tenure protections for new teachers, with six of these bills ultimately going into effect.<sup>3</sup>

### **III. Conceptual Framework**

How might adopting high stakes evaluation systems and eliminating tenure affect the decision of individuals to join the teaching profession? Evidence suggests that individuals who select into teacher preparation programs are more risk-averse and place a higher premium on job security than other college graduates (Bowen, Buck, Deck, Mills, and Shuls, 2015). Eliminating or restricting this job security could negatively affect new teacher labor supply through several mechanisms. Evidence from other labor sectors demonstrates that workers in occupations with higher unemployment risk earn higher salaries to compensate for this risk (Feinberg, 1981). Accountability reforms that weaken or eliminate job security could decrease labor supply if teacher salaries remain fixed. Reforms may also decrease new labor supply if they make the profession less enjoyable by, for example, decreasing teacher autonomy through a prescribed curriculum aligned with high-stakes tests and an increased focus on test preparation. Importantly, even if accountability reforms have no *direct* effect on job protections or satisfaction, they may still affect new labor supply if they affect the *perception* among potential entrants into the profession that teaching is a less secure or enjoyable career (Donaldson, 2016; Donaldson, Anagnostopoulos, and Yang, 2017; Kraft and Gilmour, 2016). At the same time, merit pay programs based on evaluation ratings could attract more qualified teachers into the profession, although the literature suggests that prospective teachers have little reliable information with

---

<sup>3</sup> Ohio and South Dakota passed laws eliminating tenure protections that were overturned the same year by state ballot referendums. Pennsylvania's tenure bill was vetoed by Governor Wolf. See Appendix B for further details on state laws restricting or eliminating tenure.



which to predict their future success in the classroom (Jackson, Rockoff, and Staiger, 2014; Rothstein, 2015).

We begin with a simple model of occupational choice to motivate our empirical work. Following Zarkin (1985), Brewer (1996), Falch and Strom (2005), and Dolton (2006), we assume that individuals choose between a career teaching in public schools (hereafter teaching),  $T$ , or the best alternative occupation,  $A$ , by comparing the lifetime expected utility they would receive in the two occupations. The expected utility function associated with each choice for individual  $i$ , residing in state  $s$ , is given by:

$$V_{is}^j = f(w_s^j, p_s^j, x_{is}), \quad (1)$$

where  $j \in (T, A)$ ,  $w_s^j$  and  $p_s^j$  are the state-specific pecuniary and non-pecuniary benefits, respectively, associated with a career in occupation  $j$ , and  $x_{is}$  is a vector of individual characteristics.<sup>4</sup> Let  $Y_{is}^*$  denote an unobserved latent variable measuring the difference in expected utility between teaching and the next best alternative occupation. Assuming  $Y_{is}^*$  is a linear function of  $V_{is}^T$  and  $V_{is}^A$  yields:

$$Y_{is}^* = \beta_0 + \beta_1 p_s^T + \beta_2 p_s^A + \beta_3 w_s^T + \beta_4 w_s^A + x_{is} \theta + \varepsilon_{is}, \quad (2)$$

---

<sup>4</sup> We assume an individual's expected utility comparison is based on the pecuniary and non-pecuniary benefits associated with teaching and the best alternative occupation that exists in the state in which they live when they are making their occupational choice decision. While individuals can decide to teach or pursue an alternative occupation in any state, a number of studies have found that teachers tend to seek employment in school districts that are in close proximity to the communities in which they grew up (Boyd et al. 2005; Reininger, 2011).

where  $\varepsilon_{is}$  is a random disturbance term. Let  $Y_{is}$  denote the occupation choice made by an individual such that  $Y_{is} = 1$  if  $j = T$  and  $Y_{is} = 0$  if  $j = A$ . It follows that:

$$Y_{is} = \begin{cases} 1 & \text{if } Y_{is}^* > 0 \\ 0 & \text{if } Y_{is}^* \leq 0 \end{cases} \quad (3)$$

The probability that individual  $i$ , decides on a career in teaching is then  $Prob(Y_{is} = 1) = Prob(Y_{is}^* > 0)$ . We do not observe individual decisions about whether to become a teacher. We do, however, observe the number of individuals that receive a state-specific teaching license in year  $t$ , and the number of college graduates that complete a teacher preparation program in each state and year. We therefore assume  $Prob(Y_{is} = 1)$  takes the form of a linear probability model and aggregate to the state-by-year level:

$$\sum_{i=1}^{n_{st}} Y_{ist} = n_{st}\beta_0 + n_{st}\beta_1 p_{st}^T + n_{st}\beta_2 p_{st}^A + n_{st}\beta_3 w_{st}^T + n_{st}\beta_4 w_{st}^A + \sum_{i=1}^{n_{st}} x_{ist}\theta + \sum_{i=1}^{n_{st}} \varepsilon_{ist}, \quad (4)$$

where  $n_{st}$  denotes the number of individuals making an occupational choice decision in state  $s$  and year  $t$ . Dividing equation (4) through by  $n_{st}$  yields:

$$\bar{Y}_{st} = \beta_0 + \beta_1 p_{st}^T + \beta_2 p_{st}^A + \beta_3 w_{st}^T + \beta_4 w_{st}^A + \bar{x}_{st}\theta + \bar{\varepsilon}_{st}, \quad (5)$$

where  $\bar{Y}_{st}$  is the proportion of individuals choosing teaching as their occupation.

In the empirical work that follows, we define  $n_{st}$  as the number of 18-to-65 year olds in state  $s$  in year  $t$ , which serves to scale our count outcomes by a proxy for the relative size of the

working-age pool of potential new entrants in the teaching profession in each state.<sup>5</sup> Defining  $n_s$  in terms of the *current* year provides a consistent approach for scaling our outcomes of interest and reflects that fact that prospective teachers likely make choices about where to apply for a teaching license in the same year in which they enter the labor market. While an individual's initial decision to major in education or pursue a graduate degree in education may occur between one to four years prior to entering the labor market, they can still choose to withdraw at any point in time.<sup>6</sup>

#### **IV. Data**

We conduct our analyses using an original state-by-year panel that combines measures of the number of individuals in each state planning to enter the teacher labor force, measures of state-level teacher accountability laws and time-variant state-level control variables from 2002 through 2016. Ideally, we would measure new teacher labor supply as the total number of candidates that applied for K-12 public school teaching positions for the first time. Although such a measure is unavailable at the national level, data collected by the U.S. Department of Education (DOE) on the number of initial teacher licenses granted by states each year serves as an advantageous proxy. States are required to report these data as a condition for receiving federal funding under Title II of the Higher Education Act.<sup>7</sup> These licensure count data include traditional probationary licenses as well as temporary teaching licenses such as emergency and

---

<sup>5</sup> State-specific data on the number of individuals between the ages of 18 and 65 comes from the U.S. Census.

<sup>6</sup> Our results are consistent if we used lagged measures of state populations aged 18-to-65 given the high autocorrelation of this measure within states over time.

<sup>7</sup> While states differ in the types of initial teacher licenses they issue and the requirements for these licenses, there are several elements commonly required to receive a license: 1) possess a bachelor's degree (often with a minimum grade point average); 2) complete a set of required professional courses; 3) pass state-specific licensure exams; 4) complete a minimum number of supervised student teaching or clinical hours; and 5) pass a criminal background check. Prospective teachers must enroll in a state-approved teacher preparation program (TPP) to meet these requirements.

intern teaching credentials. This measure is also robust to the rapid expansion of alternative certification programs and charter schools over the panel. Teachers who enter the profession via alternative pathways such as Teach for America are still required by states to obtain temporary licensures. The vast majority of states also require teachers who work in charter schools to obtain a state licensure.<sup>8</sup> Although licensure reciprocity agreements exist between some states, individuals seeking to teach in a new state must still acquire a state-specific initial teaching license and will be included in our data.<sup>9</sup>

We complement our licensure measure with data on the number of graduates from university-based teacher preparation programs from the IPEDS survey. We restrict our sample to bachelor's and master's degree recipients from programs in institutions of higher education preparing students for classroom teaching.<sup>10</sup> The advantage of the IPEDS data is that it disaggregates data across a number of subgroups, thus allowing us to test for heterogeneous effects across subject areas, gender, and race. These data also allow us to test for differential impacts by institutional selectivity, as measured by the Barron's index as well as the 25<sup>th</sup> and 75<sup>th</sup> percentiles of freshman SAT math scores, which have been shown to be predictive of teacher effectiveness (Rockoff, Jacob, Kane and Staiger, 2009; Jacob et al., 2016).

These advantages are offset, however, by several limitations of this measure. First, program completers may intend to teach in private schools or work outside of education. We

---

<sup>8</sup> Of the 42 states with charter school laws in 2016, 28 states (67 percent) required that teachers obtain state licenses, while 11 states (26 percent) required the majority the teachers to be licensed (~ 50 to 75 percent of teachers). Only three states (Arizona, Alabama [which only allowed charters in 2015], and Louisiana) had no licensure requirement for charter school teachers (Education Commission on the States, 2016).

<sup>9</sup> Evidence suggests that despite reciprocity agreements, costs associated with transferring licensure are quite high (National Association of State Directors of Teacher Education and Certification, N.D.; Barnum, 2017; Sindelar, Bishop, Gill, Connelly, and Rosenberg, 2007). Thus, mobility within a state is much more common than across state lines (Goldhaber, Grout, Holden, and Brown, 2015).

<sup>10</sup> See Appendix C for a detailed description of the classification of instructional programming (CIP) codes we used to identify graduates of teacher preparation programs and subject areas.

estimate that the percentage of teacher education program completers who actually go on to work in public education ranges between 46 percent (Social Science and History) and 67 percent (Special Needs), depending on subject type.<sup>11</sup> Second, program completers may seek to teach in a state other than the one in which they completed their degree. Third, this measure does not include program completers from all institutions of higher education given that only institutions that receive federal aid are required to report these data. Finally, any potential effect of teacher accountability reforms on the choice of major for undergraduate students, who constitute almost half of all teacher preparation program completers, is likely to be delayed across several years. Thus, we employ IPEDS data to facilitate exploratory analyses that are not possible with license data, while recognizing that they are noisy measures of new teacher labor supply and will likely understate any effect of state-specific accountability reforms.

We draw upon two systematic reviews of teacher evaluation reforms to code the timing of reform efforts across states (Steinberg and Donaldson, 2016; NCTQ, 2016). *Implement* *Evaluation* is coded as one in the fall of the academic year in which districts across the state first fully implemented the redesigned teacher evaluation system. Figure 1 displays the timing of full system implementation across states over time. This tractable approach, however, cannot capture important differences in system design features. States adopted a wide range of weights they assigned to test-score based measures of teacher performance ranging from 0 to 50 percent. Approximately half of the states also phased in requirements to use test-score based measures across several years following statewide implementation. Some states also delayed the use of evaluation scores to inform high-stakes personnel decisions until after initial statewide

---

<sup>11</sup> Calculations are based on employment data from the American Community Survey 3-year file from 2011-2013. These data include employment and earnings for individuals aged 25-64. We multiplied the national percentage of individuals with a given major (e.g. elementary education) who were employed full time as an elementary or secondary classroom teacher by the percentage of individuals with that degree who were found to be working at all.

implementation. We conduct several tests examining the effects of these differences across states and discuss the implications of these staggered implementation patterns in more detail below.

We compiled data on teacher tenure reforms from reviews of state statutes, case law, and prior literature.<sup>12</sup> To capture the effect of the elimination of tenure on teacher labor supply, we created an indicator variable, *Eliminate Tenure*, which is equal to one in states and years during which tenure did not exist and zero when tenure existed. We focus our coding on the six states in which tenure reform legislation was not overturned in the same calendar year by referendum or veto. Figure 2 depicts the timing of tenure reforms across states over time.

In our preferred models, we include a parsimonious set of control variables to capture pecuniary and non-pecuniary benefits associated with teaching and other occupations, namely: 1) real average teacher salaries; 2) real current per-pupil expenditures; 3) the fraction of the population age 5-17 living in families at or below the federal poverty line; 4) the percentage of the population age 5-17 that is White, Black or Hispanic; and 5) real average hourly wages in the private sector. We also include two additional variables designed to capture differences in economic conditions across states and years that may be correlated with the timing of adoption of accountability reforms and the demand for teachers, namely, annual unemployment rates and real state tax revenue per capita.<sup>13</sup>

---

<sup>12</sup> Sources include: the National Council on Teacher Quality's (NCTQ) State Teacher Policy Yearbooks, The Education Commission of the States' State Policy Database, Students First: State Policy Report Cards, and the Fordham Foundation's (2012) report entitled How Strong are Teachers' Unions? A State-by-State Comparison. We resolved discrepancies between these sources using information from states' department of education websites, Race To The Top federal grant applications, Lexis-Nexis searches of state and local newspapers, and conversations with academics and state Department of Education officials.

<sup>13</sup> Data on average hourly wages in the private sector was constructed using annual data from the CPS. Data on state unemployment rates is from the Bureau of Labor Statistics. Data on average teacher salaries and current per-pupil expenditures comes from the National Center for Education Statistics (NCES), while poverty, racial and ethnic data for the share of the population age 5-17 and state tax revenue data come from the U.S. Census. We link school data to our panel using the spring year of the academic year (e.g. 2015 for AY 2014-15). Hourly wages, teacher salaries and current per-pupil expenditures are deflated to real 2014 dollars using the consumer price index.

Table 1 presents the means and standard deviations of outcome variables used in our analysis. Across our 15-year panel, the average annual number of licenses granted per state is 15 per 10,000 adults age 18-to-65. We discuss the additional outcome variables used in exploratory analyses below.

## **V. Empirical Framework**

We estimate the effects of teacher accountability on the decision of individuals to enter the teacher labor market using a difference-in-differences framework. Our identification strategy compares changes within treated states over time to other non-treated states in the same geographic regions. This approach relies on two key assumptions, namely, that the timing of accountability reforms is as good as random and that there are no other concurrent changes in labor market conditions or teacher policies across treated states but not untreated states. Importantly, the differential timing of accountability reforms across treated states allows us to remove any regional trends in teacher labor supply and demand that might confound our estimates. Teacher accountability reforms were not the only education policy reforms that occurred during our panel period. Some states also adopted related reforms such as reducing the scope of collective bargaining rights, passing right-to-work laws, changing the length of probationary period, increasing teachers' contributions to pension plans, introducing new teacher licensure exams, adopting common core standards, and administering new common-core aligned tests. However, none of these reforms were implemented consistently in the same states and at the same time as teacher accountability reforms. In robustness tests presented below, we find that our results are largely unchanged when controlling for these additional policy reforms.

We begin with a non-parametric event-study specification. This allows us to model any anticipatory effects of the policy in a fully flexible way, as well as non-linear changes in the post-period:

$$Y_{st} = \sum_{\tau=-7}^3 \lambda I_{\tau} (year_t - \tau_0) + \Gamma X_{st} + \pi_s + \gamma_{gt} + \epsilon_{st}, \quad (6)$$

where  $Y_{st}$  is an outcome of interest for state  $s$  in year  $t$ ,  $\pi_s$  and  $\gamma_{gt}$  are state and region-by-year fixed effects, respectively, that account for fixed differences in the public teacher labor markets across states and regional labor market shocks across time, and  $\epsilon_{st}$  is a random disturbance term. Our primary outcomes of interest,  $Y_{st}$ , are ratios of the number of new teaching licenses or teacher preparation program completers per every 10,000 individuals aged 18-to-65 in a given state and year. Because equation (6) is a grouped linear probability model that is motivated by equation (5) from our conceptual framework, the variance of the random disturbance term is proportional to the denominator of our outcome, namely, the number of individuals aged 18-to-65. To improve precision, we weight all models by this denominator. We test the sensitivity of our findings to alternative weights and functional forms below and find that our results are consistent across model specifications.

The key variables of interest in (6) are the set of indicators for the years pre- and post-policy reform. The term  $(year_t - \tau_0)$  is a re-centered linear time trend for states that adopted the focal policy reform in time  $\tau_0$ , while  $I_{\tau}$  is an indicator variable that equals 1 in year  $\tau$  and 0 otherwise. Thus,  $year_t - \tau_0$  takes on a value of -1 in the year prior to the policy reform and 0 in the first year of the reform. We model relative time as a set of binary indicators for the years prior to a policy change (-7 or more to -1) as well as post-policy change (0 to 3 or more) to



capture year-specific effects on the outcome of interest.<sup>14</sup> We show results from specifications without controls as well as specifications that include the parsimonious set of covariates,  $X_{st}$ , described above. It is possible that teacher accountability reforms both directly and indirectly affected covariates such as per-pupil expenditures, teacher salaries and the composition of the student population. Adding these controls may serve to attenuate our estimates to the extent that any effects on teacher labor supply are mediated through changes to these contexts and benefits of the teaching profession. We include them in our full models as a conservative approach to account for changes in teaching contexts and benefits that were the result of other education policy reforms that may have occurred around the same time as the implementation of teacher accountability reforms.

We next adopt a standard difference-in-differences (DD) model to pool estimates across years post-reform and increase the precision of our estimates as follows:

$$Y_{st} = \beta_1 Treat_{st} + \Gamma X_{st} + \pi_s + \gamma_{tg} + \mu_{st}, \quad (7)$$

where,  $Treat_{st}$  is an indicator variable that equals unity in all years post-policy adoption,  $\mu_{st}$  is a random disturbance term and all other variables are as defined in (6). The coefficient of primary interest in (7) is  $\beta_1$ , which is the difference-in-differences estimate of the effect of a given policy reform averaged across the post-period years in our panel.

We extend the DD model in (7) to more formally test for any incremental effects and differential pre-period trends in outcomes. The effect of accountability reforms on teacher labor

---

<sup>14</sup> This event study approach is limited by the unbalanced sample of treated states in the years after accountability reforms were adopted. We observe at least four years of data post reforms for only 21 of the 44 states that adopted evaluation reforms and for only four of the six states that enacted tenure reforms.

supply may not be immediate given that the majority of states phased in system design features and/or requirements to use evaluation scores to inform high-stakes personnel decisions across several years following statewide implementation. Furthermore, it could take time for the implications of the policy to affect prospective teachers' credential decisions. We follow LaFortune, Rothstein, and Schanzenbach (forthcoming) and add two linear time trends as follows:

$$Y_{st} = \delta_1 Treat_{st} + \delta_2 Treat_{st} * (year_t - \tau_0) + \delta_3 (year_t - \tau_0) + \Gamma X_{st} + \pi_s + \gamma_{tg} + v_{st}, (8)$$

where  $Treat_{st} * (year_t - \tau_0)$  is the interaction between our relative time trend and the treatment indicator,  $v_{st}$ , is a random disturbance term and all other variables are as defined in (7). The interaction term in (8) allows for the relative time trends among treated states to differ pre- and post-reform. The coefficient on the main effect of treatment,  $\delta_1$ , captures the immediate response of new teacher labor supply in the first year of the policy change, while the coefficient associated with  $Treat_{st} * (year_t - \tau_0)$ ,  $\delta_2$ , captures any deviation from the linear trend in labor supply in the post-reform period among treated states. The coefficient associated with the relative year term  $(year_t - \tau_0)$ ,  $\delta_3$ , tests for any differential linear trends in the pre-reform period among states that adopted teacher accountability reforms relative to those that did not. This serves as a direct test for differential trends in the pre-period. We estimate both standard errors clustered at the state level as well as wild clustered bootstrap confidence intervals following Cameron, Gelbach, and Miller (2008) to account for the small number of treatment clusters for tenure reforms. In tables, we report standard errors based on the more traditional clustering at the state level as this approach produces more conservative estimates in our data.

## **VI. Results**

### *A. Descriptive Trends in Teacher Labor Supply*

National trends in the supply of potential job candidates for public teaching positions have changed substantially over the past fifteen years. Figure 3 plots trends in new teacher labor supply as measured by licenses and degree completers juxtaposed with the number of full-time equivalent teachers (FTE) employed in publicly-funded schools (traditional and charter) and private schools. For licenses, relative supply increased from 2002 to 2007 and then declined sharply during the Great Recession. The number of degree completers follows a very similar pattern with a three- to four-year lag likely reflecting that over half of all degree completers are in 4-year bachelor's degree programs. Relative to pre-recession levels in 2007, new teacher labor supply has declined by 23.4 percent as measured by licensures and 20.2 percent as measured by teacher preparation program graduates. At the same time, the size of the teacher labor market in publicly-funded schools has increased since 2002 reaching a high in 2009, falling by approximately 3.5 percentage points the following year largely due to layoffs and incentivized early retirements under the Great Recession (Kraft, 2015), and then steadily expanding through 2016. The charter sector alone has grown by over 500 percent since 2002, but remains at only 4.5 percent of the FTE teachers employed in publicly-funded schools.

The failure of new teacher labor supply to recover, even partially, to pre-recession levels stands in contrast to the steady rise in the teacher workforce in publicly-funded schools.

Although a true measure of national demand for licensed teachers does not exist, recent increases in total FTE across traditional and charter schools suggest decreased demand is unlikely to be a primary explanation for the sustained contraction. The lack of recovery also cannot be accounted

for by an expansion of the private sector where state licenses, or even formal teacher training, are not required. Steep declines in the size of the private school teacher workforce, over 15 percent since 2010, suggest decreasing demand for new teachers. We examine the degree to which teacher accountability reforms enacted by states almost entirely during the post-Recession period have played a role in the sustained contraction of the new teacher labor market.

*B. Effect of High-Stakes Teacher Evaluation on New Teacher Labor Supply*

We begin with estimates from our fully-flexible event-study model with covariates shown in Figure 4 and reported in column 2 of Appendix Table A1. Estimates should be interpreted as Intent-to-Treat (ITT) estimates of the effect of high-stakes reforms given the wide variability in implementation across and within states (Kraft and Gilmour, 2017). These ITT estimates serve to answer the relevant question from a policymaker’s perspective—what is the effect of passing and implementing a state-wide high-stakes evaluation reform? The parameter estimates with associated 95 percent confidence intervals for the years pre- (hollow dot) and post-reform (solid dot) are strongly suggestive of a negative effect of evaluation reforms on new teacher labor supply. The number of teacher licenses granted remains largely unchanged in pre-period years and then declines steadily starting the year in which high-stakes evaluation reforms are implemented statewide. These individual point estimates become statistically significant in the second year and suggest that the effects of the reform increase in post-reform years.

Results from our standard DD model given by (7) confirm the effect of high-stakes evaluation reforms. In Table 2, we estimate that high-stakes evaluation reforms reduced licenses granted in a state by 2.51 per 10,000 18-to-65 year olds, on average, in our specification that includes controls. This represents a 17 percent reduction in the average number of licenses granted in the post-policy reform years among treated states, relative to the average number

licenses granted by states across the years in our panel. Estimates based on our parametric DD specification that includes linear trends (equation (8)) also confirm the patterns suggested by our event study analyses. In column 4 of Table 2, we find a negative initial shock to new teacher supply, as well as a significant downward linear trend in supply among treated states in the post-policy reform years of -1.34 licenses per 10,000. Consistent with the visual evidence in the event study, we also fail to reject the null hypothesis of no differential pre-reform trends for states that adopted evaluation reforms: the coefficient on the pre-period trend variable (*Year\_R*) is both small in magnitude (0.33) and statistically insignificant. For both our standard and linear DD models, comparing results between our baseline and preferred specifications illustrates that adding our set of time-varying controls has little effect on the estimates.

We next explore potential heterogeneity in the effects of evaluation reform on new teacher labor supply. One possible source of heterogeneity is differences in the strength of state teacher unions, who strongly opposed changes to the evaluation system. In states with strong unions, lobbying could have led to more modest changes to teacher evaluation systems. Other potential sources of heterogeneity include differences in the degree of difficulty for dismissing ineffective teachers prior to reforms and differences in the use of test-based evaluation measures as part of the new teacher evaluation systems. To examine these possibilities, we extend the results reported in Table 2 by testing for potential moderation effects by union strength, obstacles to dismissing ineffective teachers, and the degree to which evaluation reforms incorporate test-based evaluation measures. We describe the specific measures we use in Appendix D and report our result in Appendix Table D1. Across all models, we find no evidence of moderation effects. The estimated coefficients on the additional interaction terms are small in magnitude and insignificant across all specifications. These findings suggest that differences in union strength,

pre-reform conditions and the use of test-based performance measures do not appear to be driving our results.

### *C. Effect of Eliminating Tenure on New Teacher Labor Supply*

Figure 5 presents estimates of the effect of restricting or eliminating teacher tenure on the number of teaching licenses awarded using the fully-flexible event-study estimates reported in column 4 of Appendix Table A1. In contrast to our estimates for high-stakes evaluation reforms, these estimates can be interpreted as capturing the Average Treatment Effect (ATE) of repealing tenure. We again see that the number of teacher licenses granted remains largely unchanged in pre-period years among the states that adopted tenure reforms. In the post-period, the point estimates are uniformly negative but statistically significant only in year 2, providing suggestive evidence of both the validity of our difference-in-differences design and the negative effect of tenure reforms on the number of licenses granted.

Table 3 presents estimates of the effect of eliminating tenure based on our parametric DD specifications given by equations (7) and (8). We estimate that repealing tenure caused a decrease of 2.82 licensures per 10,000 18-to-65 year olds, using our standard DD specification including controls. Allowing for linear trends pre- and post-reform among treated states reinforces these findings. Results from this linear DD specification suggest the impact of tenure reforms was more immediate than the effect of high-stakes evaluation reforms but diminished over time. Importantly, we once again fail to reject the null hypothesis of a pre-reform trend that was unique to states that adopted tenure reforms. The coefficient on *Year\_R* is near zero (-0.19) and statistically insignificant. Comparing estimates across models with and without controls again demonstrates the robustness of our findings.

Given the small number of treated states, we further examine the degree to which any one state is driving our results by individually omitting each of the six states and re-estimating our primary DD models. Appendix Table A2 presents estimates across these six subsamples. The estimates are relatively stable, suggesting that our results are not the product of a single outlier. Removing Florida from the sample results in the largest estimated treatment effect of tenure reforms from our standard DD models at -4.06, while removing North Carolina results in the smallest, -1.77, which is no longer statistically significant. Instead, effects in the sample without North Carolina appear in our linear DD models with a large immediate shift downward in teacher licensures, an effect that is offset over time with a positive linear trend. Corresponding linear DD estimates across the leave-one-out samples consistently show a large negative intercept shift downward. Overall, these results confirm the negative relationship between tenure reforms and new teacher labor supply but suggest the specific time dynamic of this effect may differ across states.

#### *D. Simultaneous Treatments*

In many instances, state legislatures packaged multiple teacher accountability initiatives into a single bill or legislative session. For example, in Florida and Idaho, legislatures eliminated teacher tenure in 2011, the same year as the establishment of new high-stakes teacher evaluation systems. Similarly, Kansas and Louisiana restricted tenure in the same year that their new high-stakes teacher evaluation systems were first implemented statewide. We examine the degree to which each of these two reforms are independently responsible for depressing new labor supply by estimating specifications that simultaneously account for the effect of evaluation and tenure reforms on teacher labor supply and by restricting the analytic sample to only those states that did not pass tenure reforms.

In Table 4, we compare the results from our DD models, where indicators for evaluation and tenure reforms were included separately, to estimates where they are modeled simultaneously. Starting with our standard DD estimates, we find that the effect of high-stakes evaluation reforms, conditional on tenure reforms, remains large and statistically significant. Our point estimate is slightly attenuated from -2.51 to -2.25. As a result, the implied percent reduction in the mean number of licenses granted falls from 17 percent in column 1 to 15 percent in column 3. Point estimates for the effect of tenure, conditional on evaluation reforms also fall from -2.82 to -2.34, but are no longer statistically significant. Estimates from linear DD model with evaluation and tenure reforms modeled simultaneously (column 7) confirm the dynamic and independent effects of both reforms. The effect of evaluation is both a moderate initial downward shift followed by a continued downward linear trend, while the effect of tenure is concentrated in an immediate downward shift in new labor supply. All of these associated terms are significant in the joint model.

Evaluation and tenure reforms were not, however, the only education policy reforms pursued by states during this time period. Other important reforms included restricting or eliminating mandatory collective bargaining (6 states), eliminating mandatory agency fees (i.e. “right-to-work”) (3 states), requiring new teachers to pass a basic skills test for certification (27 states), requiring new teachers to pass an exam testing professional knowledge (34 states), requiring new teachers to pass content tests in their subject areas (24 states), adopting Common Core State Standards (CCSS) (45 states), administering new Common Core (CC) aligned tests (29 states), and increasing teacher’s mandatory contributions to pension funds (32 states).<sup>15</sup> For example, Tennessee passed laws that substantially restricted the scope of mandatory collective

---

<sup>15</sup> We describe these measures and their data sources in detail in Appendix E.



bargaining in 2011, the same year in which the state implemented its new high-stakes teacher evaluation system. It is possible that these policy changes, many of which were concurrent with evaluation and tenure reforms, are conflated with our estimates.<sup>16</sup> In columns 4 and 8 of Table 4, we report results from joint models that also control for the reforms described above: collective bargaining, right-to-work, licensure exams, CCSS, CC aligned tests, and pension reforms. Controlling for these additional policy reforms has little effect on our results, confirming that our accountability estimates are not confounded by these concurrent policy reforms.

As a further test of the independent effect of evaluation on new teacher labor supply, we re-estimate the specifications reported in Table 2 based on a restricted sample that excludes the six states that repealed tenure. These results, presented in Table A3, further reinforce the independent effect of evaluation reforms. We find a positive and significant effect of evaluation reforms of a 14 percent decline in licenses using our standard DD model with controls.

## **VII. Differential Effects by Subject, Institutional Selectivity, Race and Gender**

We extend our primary analyses using the number of university-based teacher preparation program degree completers to allow for a range of sub-group analyses. As expected, average estimates for both accountability measures are negative but small and imprecise, given the delayed response of this measure and our short post-period window. We report formal estimates from our event study models in Appendix Table A1 and from standard and linear DD models in Table 5. Point estimates from our standard DD model suggest that evaluation reforms reduced the number of degree completers by 5 percent and that tenure reforms reduced the

---

<sup>16</sup> As a further robustness check, we fit models controlling for the number of years of consecutive experience teachers must have to be eligible for tenure (probationary period) and find no difference in our estimates for evaluation reforms. We do not include this measure in Table 4 because it is undefined for states that have eliminated tenure. Results available on request.

number of degree completers by 8 percent, although these estimates are too imprecise to reject the possibility of a null effect.

We next test for evidence of possible differential effects by teacher subject area. Changes in new teacher supply by subject area vary substantially across our 15-year panel. Figure 6 presents national trends in the number of degree completers by subject area relative to 2002. The supply of graduates from math education and special education degree programs increased 42.5 and 29.4 percent over this period. In contrast, the number of elementary, science, social studies and English education degree completers has decreased by 41.0, 38.5, 24.4 and 9.1 percent, respectively. With the exception of prospective science teachers, these patterns suggest an overall pattern of market corrections with supply increasing in hard-to-staff subject areas and decreasing in subject areas with excess supply. Despite these encouraging trends, teacher shortages remain a real challenge in some subject areas. As shown in Figure 7, for the 2017-18 school year, 88 percent of states designated math as a shortage area, followed by 80 percent for special education and 74 percent for science. This contrasts with fewer than half the states that designated English, elementary, and social studies as shortage areas.

We begin by testing for heterogeneous effects of accountability reforms on shortage areas (math, special education, science) vs. non-shortage areas (elementary, English, and social studies). As shown in Table 6, we find no evidence of differential effects of accountability across shortage and non-shortage subject areas. The relative magnitude of estimates for shortage and non-shortage licensure areas from our standard models are 1 and 3 percent declines due to evaluation and 9 and 9 percent decline for tenure. Analyses using subject-specific outcomes point to consistently larger negative effects on English teachers (23 and 24 percent declines for

evaluation and tenure reforms) and science teachers (33 and 15 percent declines for evaluation and tenure reforms), consistent with overall trends in Figure 6.

We next explore the potential effects accountability reforms had on prospective teacher quality by testing for post-reform changes in the selectivity (as measured by an institution's Barron's ranking and average freshman SAT scores) of the institutions where prospective teachers completed their teacher training. In Table 7, we report estimates of the effect of accountability reforms on the number of degree completers from very competitive colleges, competitive colleges, less competitive colleges, and unranked colleges according to Barron's ranking system.<sup>17</sup> In Table 8, we present estimates of changes in the 25<sup>th</sup> and 75<sup>th</sup> percentiles of freshman SAT scores in math for the institution in which graduates completed their teacher preparation program. For evaluation reforms, patterns in the point estimates by college selectivity are inconsistent with the largest negative effects concentrated among prospective students at unranked schools, a decline of 41 percent. At the same time, we find negatively signed and statistically insignificant point estimates for the effect of evaluation reforms on average freshman SAT math scores. Estimates across both measures of college selectivity suggest tenure reforms may have had positive effects on the qualifications of new teacher supply. The pattern of results across graduates based on Barron's rankings suggests a larger negative effect among less competitive and unranked colleges and the smallest effect for very competitive colleges. We also find significant positive effects on trends in the 25<sup>th</sup> and 75<sup>th</sup> percentile of average freshman SAT math scores at institutions where students attended teacher preparation programs in the years post-tenure reforms.

---

<sup>17</sup> While these measures of university selectivity are consistent across schools, they are based on undergraduate admissions statistics and thus may be weak proxies for the quality of graduate teacher preparation programs.

Finally, we test for differential effects by gender and race and report the results in Table 9. Estimates are remarkably consistently across both gender and race for the effects of evaluation reforms. We find little difference in effects by gender for tenure reforms. Effects of tenure reforms by race are suggestive of a disproportionately large negative effect on the supply of Black teachers. Based on our standard model, we estimate a 44 percent reduction in the number of Black graduates from teacher preparation programs relative to a 9 percent reduction for Hispanic and White graduates. This point estimate is large, but is only marginally significant, and we find no corresponding evidence of differential effects on the supply of Black teachers due to evaluation reforms.

## **VIII. Robustness Tests & Extensions**

### *A. Falsification Tests*

The effect of teacher accountability reforms on new labor supply should be isolated to future professionals intending to enter the teacher labor market in publicly-funded schools. As a falsification test, we estimate impacts on a proxy measure for potential new labor supply in a private sector industry that also requires state certification. Specifically, we use the number of individuals who take the Certified Public Accountant (CPA) exam for the first time in a given state and year.<sup>18</sup> Similar to our licensure data, we scale this outcome per 10,000 18-to-65 year olds. If broader state-level economic factors correlated with the timing of teacher accountability reforms are driving our results, we would expect to find similar negative effects of these policy

---

<sup>18</sup> The American Institute of Certified Public Accountants (AICPA) prepares, administers, and scores the CPA exam in all U.S. states and territories. Requirements for CPA candidates are similar to those for teaching candidates – while all CPA candidates are required to have *at least* a bachelor’s degree and complete a set number of accounting courses, the CPA examination/licensure requirements vary from state-to-state. The state-by-year cohorts are calculated by the National Association of State Boards of Accountancy (NASBA) in partnership with AICPA. All candidates who begin the CPA examination process (begin any one of the four sub-exams) in the same calendar year are members of a state-year cohort.

reforms on CPA exam takers. The results reported in Appendix Table A4 reveal little evidence that would suggest our findings are picking up general trends in the college-educated labor market in states that adopted teacher accountability reforms. We find no significant effect of either evaluation or tenure reforms on the number of CPA exam takers in our standard or linear DD models, with positively signed point estimates.

### *B. Endogenous Spillover*

Our identification strategy compares changes within treated states over time to other non-treated states in the same geographic regions. If the introduction of teacher accountability reforms caused prospective teaching candidates to seek teaching positions in nearby states, then our estimates would overstate the effect of these reforms. Comparison states that share borders with treated states would experience a concurrent positive treatment effect inflating the treatment-control contrast artificially. In practice, existing evidence suggests such cross-state mobility in the teacher labor market is unlikely. Studies find that teacher employment preferences are extremely localized (Boyd, Lankford, Loeb, and Wyckoff, 2005; Reininger, 2012) and that cross-state mobility is rare (Goldhaber, Grout, Holden and Brown, 2015). However, pre-service teachers may be more willing to cross state lines than teachers already established in a job where state-specific licensure regulations, seniority rules, and pension structures can impose substantial costs on mobility.

We formally test for endogenous spillover by re-estimating our models with an indicator for states that share borders with treated states in the years during which contiguous states were treated. If a state in this spillover set later adopts an accountability reform, it is recoded as zero for all years it is treated. We narrow our analytic window for evaluation reforms to examine treatment spillover effects through 2012 to focus on the first nine states that implemented

evaluation reforms statewide. As shown in Figure 1, the frequency of evaluation reforms accelerated in 2013 with a geographically diverse set of 13 states. By 2013, nearly every non-treated state becomes part of our spillover comparison groups, limiting the usefulness of this test.

Results of our tests for endogenous spillover are reported in Appendix Table A5. In this restricted panel, we find similar negative and statistically significant effects of evaluation and tenure reforms on licensures. We also find little evidence of endogenous spillover among contiguous non-treated states. For both evaluation and tenure, we fail to reject the null hypothesis of no spillover effects on bordering states.

### *C. Alternative Specifications*

In Appendix Table A6, we examine the sensitivity of the main results reported in Tables 2 and 4 to our choice of weights and functional form. We report results from three alternative specifications for both our standard and linear DD models. First, we replace our scaling and weighting variable, the number of 18-to-65 year olds per 10,000 in a state-year cell, with an alternative measure, the number of 22-to-25 year olds per 1,000 in a state-year cell. This allows us to focus our estimates relative to the population of recent college graduates who compose the majority of new entrants into the teacher labor market. Second, we use our original scaling factor of the number of 18-to-65 year olds per 10,000 but no longer weight our models using this measure. Third, we specify our count outcomes as logs without any scaling factor and control for the log number of 18-to 65 year olds per 10,000. Results are quite robust across specifications. For our standard model with controls, these alternative specifications produce estimates of an 18 percent, 16 percent and 15 percent decline in new labor supply caused by evaluation reforms, compared to our main estimate of 17 percent. Alternative specifications for tenure reforms result in declines of 21 percent, 18 percent, and 20 percent, relative to our main estimate of 19 percent.

#### *D. Effects of Teacher Wages*

We examine whether an increase in wages potentially offset the real or perceived loss of job security in states that adopted teacher accountability reforms using two different measures. The first is the control variable we use in our models, real average public school teacher wages calculated using district reported total FTE instructional staff salaries collected by NCES. Second, we use average total real wages for public school teachers from the 2005 – 2016 Public Use Microdata Sample (PUMS) of the American Community Survey (ACS), adjusted for age and educational attainment. This adjustment serves to account for differences in average wages caused by changes in the distribution of experience and educational attainment among public school teachers rather than changes in the underlying salary schedules.

We report the effects of evaluation reforms on both wage measures in Appendix Table A6 Panel A and the effect of tenure reforms in Panel B. We find no evidence of any compensating differentials that might offset the loss in job security (perceived or actual) due to teacher accountability reforms. Estimates across all specifications are small in magnitude (less than \$450), negatively signed and statistically insignificant with the exception of one marginally significant result. This is consistent with our primary finding of a large negative effect of teacher accountability reforms on the number of new licenses granted. Together, these findings illustrate that there is little evidence to suggest that wage effects are driving the contraction in the new labor supply we observe or serving to offset this contraction in any way.

### **IX. Conclusion**

Education policy over the past decade has focused, in large part, on improving human capital in schools through teacher labor market reforms. Many of these policies have identified

teacher accountability as a promising mechanism for raising teacher quality. These reform efforts, and the research base upon which they draw, have largely failed to consider the potential consequences of accountability reforms on the future supply of new teachers. Our findings document how both adopting high-stakes evaluation systems and eliminating tenure protections reduce the supply of new teaching candidates available to public schools. Flexible models also reveal dynamic labor market responses to these reforms over time.

Evaluation reforms resulted in a gradual but persistent decline in new teacher labor supply, possibly due to the staggered implementation of high-stakes evaluation systems and the delayed use of evaluation ratings for high-stakes decisions in many states. In contrast, the immediate contraction of supply after states repealed tenure was followed by a gradual return to pre-reform levels, suggesting that prospective teachers may have updated their beliefs about the threat posed by the loss of job protections when they did not observe districts acting on this threat.

These results also provide new evidence of the competitiveness of the teacher labor market. The loss of non-pecuniary benefits due to accountability reforms was not offset by any compensatory changes in teacher salaries. New entrants into the labor market who might otherwise have become teachers chose not to enter this labor sector. A contraction in the supply of new teachers could have potential benefits, if it helped correct the large and persistent imbalances in supply across subject areas. It might also serve to increase the average quality of new teachers if less qualified and capable candidates were discouraged from entering the profession. While the data available to examine these questions are more limited, we find no evidence that accountability reforms served to correct imbalances in supply and demand across subject areas. We find some suggestive evidence that tenure reforms induced less qualified



teaching candidates to seek employment outside of education. At the same time, we find suggestive evidence that tenure reforms may have substantially reduced the supply of Black teachers. This is particularly concerning given efforts to diversify the teacher workforce in many states and evidence of the positive effects for Black students of being taught by Black teachers (Dee, 2004; Dee 2005; Gershenson, Holt, and Papageorge, 2016; Gershenson et al., 2017).

Enrollment in K-12 public schools in the United States is projected to increase by over a million students in the next decade, an increase of 2 percent from current levels. Even with new personalized learning technologies, it is hard to imagine a future where demand for classroom teachers is not increasing given expanding enrollments and a labor force where 29 percent of teachers are over the age of 50 (NCES, 2017). Understanding the consequences of education policy reforms on teacher labor supply will remain a key element of efforts to improve human capital in the education sector. This is particularly important for subject areas with teacher shortages, as well as for the pipeline of new teachers needed to work in hard-to-staff schools in rural and low-income settings.

## References

- Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and student achievement in the Chicago public high schools." *Journal of Labor Economics* 25, no. 1: 95-135. <https://doi.org/10.1086/508733>
- Barnum, Matt. 2017, February 15. "The Certification Maze: Why Teachers Who Cross State Lines Can't Find Their Way Back to the Classroom" The 74 Media.
- 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, no. 1: 113-132. <https://doi.org/10.1002/pam.20072>
- Bacolod, Marigee P. 2007. "Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality." *The Review of Economics and Statistics* 89, no. 4: 737-751. <https://doi.org/10.1162/rest.89.4.737>
- Boskin, Michael J. 1974. "A conditional logit model of occupational choice." *Journal of Political Economy* 82, no. 2, Part 1: 389-398. <https://doi.org/10.1086/260198>
- Bowen, Daniel H., Stuart Buck, Cary Deck, Jonathan N. Mills, and James V. Shuls. 2015. "Risky business: an analysis of teacher risk preferences." *Education Economics* 23, no. 4: 470-480.
- Brewer, Dominic J. 1996. "Career paths and quit decisions: Evidence from teaching." *Journal of Labor Economics* 14, no. 2: 313-339. <https://doi.org/10.1086/209813>
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based Improvements for inference with clustered errors." *The Review of Economics and Statistics* 90, no. 3: 414-427.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *The American Economic Review* 104, no. 9: 2633-2679. <https://doi.org/10.1257/aer.104.9.2633>
- Cullen, Julie Berry, Cory Koedel, and Eric Parsons. 2016. *The Compositional Effect of Rigorous Teacher Evaluation on Workforce Quality*. No. w22805. National Bureau of Economic Research, <http://www.nber.org/papers/w22805.pdf>
- Dee, Thomas S., and James Wyckoff. 2015. "Incentives, selection, and teacher performance: Evidence from IMPACT." *Journal of Policy Analysis and Management* 34, no. 2: 267-297. doi:10.1002/pam.21818
- Dee, Thomas S. 2004. "Teachers, race, and student achievement in a randomized experiment." *The Review of Economics and Statistics* 86, no. 1: 195-210.
- Dee, T. S. 2005. A teacher like me: Does race, ethnicity, or gender matter?. *The American*

*economic review*, 95(2), 158-165.

Dolton, P.J. 2006. "Teacher Supply." In E. Hanushek and F. Welch (eds) *Handbook of the Economics of Education*. Amsterdam: North-Holland, 1079-1161.

Donaldson, Morgaen L. 2009. "So long, Lake Wobegon? Using teacher evaluation to raise teacher quality." *Center for American Progress*.

Donaldson, Morgaen L. 2016. "Teacher Evaluation Reform: Focus, Feedback, and Fear." *Educational Leadership* 73, no. 8: 72-76.

Donaldson, Morgaen L., and John P. Papay. 2015. "Teacher evaluation for accountability and development." *Handbook of research in education finance and policy*: 174-193.

Donaldson, Morgaen, Dorothea Anagnostopoulos, and Min Yang. 2017. "How do teachers respond to teacher evaluation? The role of emotions." University of Connecticut Working Paper.

Education Commission of the States. 2016. "Charter Schools - Do teachers in a charter school have to be certified?" Charter Schools - Do teachers in a charter school have to be certified? January 2016. <http://ecs.force.com/mbdata/mbquestNB2?rep=CS1525>

Falch, Torberg, and Bjarne Strøm. 2005. "Teacher turnover and non-pecuniary factors." *Economics of Education Review* 24, no. 6: 611-631.  
<https://doi.org/10.1016/j.econedurev.2004.09.005>

Feinberg, Robert M. 1981. "Earnings-risk as a compensating differential." *Southern Economic Journal*: 156-163. <https://doi.org/10.2307/1058607>

Fullan, Michael. 2011. *Choosing the wrong drivers for whole system reform*. Melbourne: Centre for Strategic Education. <http://ccee-ca.org/documents/CCEE%20Local%20Control%20and%20Continuous%20Improvement%20Workshop%20Handout.pdf>

Gershenson, Seth, Cassandra Hart, Constance Lindsay, and Nicholas W. Papageorge. 2017. "The long-run impacts of same-race teachers." *IZA Discussion Paper No. 10630*

Gershenson, Seth, Stephen B. Holt, and Nicholas W. Papageorge. 2016. "Who believes in me? The effect of student-teacher demographic match on teacher expectations." *Economics of Education Review* 52: 209-224.

Goldhaber, Dan, and Michael Hansen. 2010. "Using performance on the job to inform teacher tenure decisions." *The American Economic Review* 100, no. 2: 250-255.  
<http://www.jstor.org/stable/27804999>

Goldhaber, Dan, Cyrus Grout, Kristian L. Holden, and Nate Brown. 2015. "Crossing the border? Exploring the cross-state mobility of the teacher workforce." *Educational Researcher* 44, no. 8: 421-431. <https://doi.org/10.3102/0013189X15613981>

Gordon, Robert James, Thomas J. Kane, and Douglas Staiger. 2006. *Identifying effective teachers using performance on the job*. Washington, DC: Brookings Institution.

Hansen, Michael. 2009. "How Career Concerns Influence Public Workers' Effort: Evidence from the Teacher Labor Market. Working Paper 40." *National Center for Analysis of Longitudinal Data in Education Research*. <http://files.eric.ed.gov/fulltext/ED509686.pdf>

Goldhaber, Dan, and Jane Hannaway. 2009. *Creating a New Teaching Profession*. Urban Institute Press. 2100 M Street NW, Washington, DC 20037.

Jacob, Brian A. 2013. "The effect of employment protection on teacher effort." *Journal of Labor Economics* 31, no. 4: 727-761. <https://doi.org/10.1086/669942>

Jacob, Brian, Jonah E. Rockoff, Eric S. Taylor, Benjamin Lindy, and Rachel Rosen. 2016. *Teacher applicant hiring and teacher performance: Evidence from DC public schools*. No. w22054. National Bureau of Economic Research.

Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger. 2014. "Teacher effects and teacher-related policies." *Annu. Rev. Econ.* 6, no. 1: 801-825. <https://doi.org/10.1146/annureveconomics-080213-040845>

Jones, Michael D. 2015. "How do teachers respond to tenure?." *IZA Journal of Labor Economics* 4, no. 1: 8.

Kraft, Matthew A. 2015. "Teacher layoffs, teacher quality, and student achievement: Evidence from a discretionary layoff policy." *Education Finance and Policy*. [https://doi.org/10.1162/EDFP\\_a\\_00171](https://doi.org/10.1162/EDFP_a_00171)

Kraft, Matthew A., and Allison F. Gilmour. 2016. "Can principals promote teacher development as evaluators? A case study of principals' views and experiences." *Educational Administration Quarterly* 52, no. 5: 711-753. 46(5), 234-249.

Kraft, Matthew A., and Allison F. Gilmour. 2017. "Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher effectiveness." *Educational Researcher*, 46, no. 5: 234-249.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (forthcoming). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*.

Loeb, Susanna, Luke C. Miller, and James Wyckoff. 2015. "Performance screens for school improvement: The case of teacher tenure reform in New York City." *Educational Researcher* 44, no. 4: 199-212. <https://doi.org/10.3102/0013189X15584773>

National Association of State Directors of Teacher Education and Certification (NASDTEC). "Interstate Agreement: Facilitating Mobility of Educational Personnel." [nasdtec.site-ym.com](https://nasdtec.site-ym.com/?page=Interstate). <https://nasdtec.site-ym.com/?page=Interstate>[9/15/2016 11:52:35 AM]

National Council on Teacher Quality (NCTQ). April 2016. *State-by-State Evaluation Timeline Briefs*. Washington, DC: NCTQ.  
[http://www.nctq.org/dmsStage/Evaluation\\_Timeline\\_Brief\\_Overview](http://www.nctq.org/dmsStage/Evaluation_Timeline_Brief_Overview)

National Center for Education Statistics (NCES). 2017. "Characteristics of public elementary and secondary school teachers in the United States." United States Department of Education.  
<https://nces.ed.gov/pubs2017/2017072.pdf> (accessed 10/12/2017).

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, no. 2: 127-145. [https://doi.org/ 10.3102/0162373711420864](https://doi.org/10.3102/0162373711420864)

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2015. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (2005): 417-458.

Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *The American Economic Review* 94, no. 2: 247-252.  
<http://www.jstor.org/stable/3592891>

Rockoff, Jonah E., Brian A. Jacob, Thomas J. Kane, and Douglas O. Staiger. 2011 "Can you recognize an effective teacher when you recruit one?." *Education Finance and Policy*, 6(1): 43-74.

Rothstein, Jesse. 2014. "Teacher quality policy when supply matters." *The American Economic Review* 105, no. 1: 100-130.

Sartain, Lauren, and Matthew P. Steinberg. 2016. "Teachers' labor market responses to performance evaluation reform: Experimental evidence from Chicago public schools." *Journal of Human Resources* 51, no. 3: 615-655.

Sindelar, Paul T., Anne G. Bishop, Michele Gregoire Gill, Vincent Connelly, and Michael S. Rosenberg. 2007. "Getting teachers where they're needed most: The case for licensure reciprocity." *Teacher Education and Special Education* 30, no. 2: 103-114. [https://doi.org/ 10.1177/088840640703000205](https://doi.org/10.1177/088840640703000205)

Staiger, Douglas O., and Jonah E. Rockoff. 2010. "Searching for effective teachers with imperfect information." *The Journal of Economic Perspectives* 24, no. 3: 97-117.  
<http://www.jstor.org/stable/20799157>

Steinberg, Matthew P., and Morgaen L. Donaldson. 2016. "The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era." *Education Finance*

and Policy. [https://doi.org/10.1162/EDFP\\_a\\_00186](https://doi.org/10.1162/EDFP_a_00186)

Strunk, K. O., N. Barrett, N., and J.A. Lincove. 2017. When tenure ends: the short-run effects of the elimination of Louisiana's teacher employment protections on teacher exit and retirement. Education Research Alliance for New Orleans, Technical Report.

Thomas, E., Wingert, P., E. Conant, and S. Register. 2010. Why we can't get rid of failing teachers. *Newsweek*, 155(11), 24-27.

Tucker, Pamela D. 1997. "Lake Wobegon: Where all teachers are competent (or, have we come to terms with the problem of incompetent teachers?)." *Journal of Personnel Evaluation in Education* 11, no. 2: 103-126.

U.S. Department of Education, National Center for Education Statistics. 2016. *Digest of Education Statistics, 2015* (NCES 2016-014), [Introduction](#) and [Chapter 2](#). Via: <https://nces.ed.gov/fastfacts/display.asp?id=28>

U.S. Department of Education, Pathways to Teaching, Office of Post Secondary Education, Higher Education Act Title II Reporting System. May 19, 2016. Accessible at: [https://title2.ed.gov/Public/46608\\_Final\\_Title\\_II\\_Infographic\\_Booklet\\_Web.pdf](https://title2.ed.gov/Public/46608_Final_Title_II_Infographic_Booklet_Web.pdf)

Weisberg, Daniel, Susan Sexton, Jennifer Mulhern, David Keeling, Joan Schunck, Ann Palcisco, and Kelli Morgan. 2009. "The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness." *New Teacher Project*. <http://files.eric.ed.gov/fulltext/ED515656.pdf>

Winkler, Amber M., Janie Scull, and Dara Zeehandelaar. 2012. "How Strong Are US Teacher Unions? A State-by-State Comparison." *Thomas B. Fordham Institute*.

Winters, Marcus A., and Joshua M. Cowen. 2013a. "Who would stay, who would be dismissed? An empirical consideration of value-added teacher retention policies." *Educational Researcher* 42, no. 6: 330-337. <https://doi.org/10.3102/0013189X13496145>

Winters, Marcus A., and Joshua M. Cowen. 2013b. "Would a value-added system of retention improve the distribution of teacher quality? A Simulation of Alternative Policies." *Journal of Policy Analysis and Management* 32, no. 3: 634-654. <https://doi.org/10.1002/pam.21705>

Zabalza, Antoni. "The determinants of teacher supply. 1979. " *The Review of Economic Studies* 46, no. 1: 131-147. <http://www.jstor.org/stable/2297177>

Zarkin, Gary A. 1985. "Occupational choice: An application to the market for public school teachers." *The Quarterly Journal of Economics* 100, no. 2: 409-446. <https://doi.org/10.2307/1885389>

## Tables

Table 1: Outcome Descriptive Statistics

| Outcome                                                         | Obs | Mean     | Std. Dev. | Date Range |
|-----------------------------------------------------------------|-----|----------|-----------|------------|
| Teacher Licensures (per 10,000 18-to-65 year olds)              | 749 | 14.97    | 5.80      | 2002-2016  |
| NCES Average Salary                                             | 750 | 54999.32 | 8053.98   | 2002-2016  |
| ACS Adjusted Average Salary                                     | 600 | 46625.20 | 6577.10   | 2005-2016  |
| CPA (per 10,000 18 to 65 year olds)                             | 550 | 2.61     | 3.41      | 2006-2016  |
| Teacher Prep Program Completers (per 10,000 18-to-65 year olds) | 750 | 13.53    | 5.35      | 2002-2016  |
| Barron's Very Competitive                                       | 750 | 3.07     | 2.35      | 2002-2016  |
| Barron's Competitive                                            | 750 | 6.55     | 3.45      | 2002-2016  |
| Barron's Less Competitive                                       | 750 | 2.69     | 2.52      | 2002-2016  |
| Barron's Unranked                                               | 750 | 1.01     | 2.39      | 2002-2016  |
| BA                                                              | 750 | 7.15     | 3.71      | 2002-2016  |
| MA                                                              | 750 | 6.37     | 3.62      | 2002-2016  |
| Non-Shortage Licensure Areas                                    | 750 | 4.64     | 2.64      | 2002-2016  |
| Elementary                                                      | 750 | 3.84     | 2.32      | 2002-2017  |
| English                                                         | 740 | 0.54     | 0.48      | 2002-2018  |
| Social Studies                                                  | 702 | 0.29     | 0.29      | 2002-2019  |
| Shortage Licensures Areas                                       | 750 | 1.73     | 1.23      | 2002-2020  |
| Math                                                            | 706 | 0.22     | 0.30      | 2002-2021  |
| Science                                                         | 721 | 0.21     | 0.32      | 2002-2022  |
| Special Education                                               | 750 | 1.32     | 0.92      | 2002-2023  |
| Female                                                          | 750 | 10.59    | 4.23      | 2002-2024  |
| Male                                                            | 750 | 2.94     | 1.24      | 2002-2025  |
| Asian                                                           | 750 | 0.23     | 0.51      | 2002-2026  |
| Black                                                           | 750 | 0.75     | 0.92      | 2002-2027  |
| Hispanic                                                        | 750 | 0.50     | 0.74      | 2002-2028  |
| White                                                           | 750 | 10.04    | 4.83      | 2002-2029  |

Notes: Statistics are weighted by the number of 18-to-65 year olds in a state.

Table 2: The Effect of High-Stakes Teacher Evaluation Reforms on the Number of New Teaching Licenses

|                                 | (1)                 | (2)                 | (3)                 | (4)                 |
|---------------------------------|---------------------|---------------------|---------------------|---------------------|
| Implement Evaluation            | -2.686**<br>(0.835) | -2.505**<br>(0.861) | -2.369**<br>(0.821) | -2.366**<br>(0.832) |
| Implement Evaluation * Year_R   |                     |                     | -1.308*<br>(0.527)  | -1.341*<br>(0.571)  |
| Year_R                          |                     |                     | 0.231<br>(0.212)    | 0.330<br>(0.204)    |
| % change relative to state mean | -18%                | -17%                |                     |                     |
| Controls                        |                     | Y                   |                     | Y                   |
| n                               | 749                 | 749                 | 749                 | 749                 |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls include real average hourly wages in the private sector, unemployment rates, real state tax revenue per capita, the fraction of the population ages 5 to 17 that are White, Black, Hispanic, and living below the federal poverty line, and measures specific to K-12 public education: real average teacher salaries and real current per pupil expenditures. All models include state and region-by-year fixed effects.



Table 3: The Effect of Eliminating Tenure on the Number of New Teaching Licenses

|                                 | (1)     | (2)     | (3)       | (4)      |
|---------------------------------|---------|---------|-----------|----------|
| Eliminate Tenure                | -2.561* | -2.822+ | -4.485*** | -3.994** |
|                                 | (1.270) | (1.430) | (1.272)   | (1.242)  |
| Eliminate Tenure* Year_R        |         |         | 1.170+    | 1.217+   |
|                                 |         |         | (0.692)   | (0.614)  |
| Year_R                          |         |         | -0.063    | -0.188   |
|                                 |         |         | (0.388)   | (0.384)  |
| % change relative to state mean | -17%    | -19%    |           |          |
| Controls                        |         | Y       |           | Y        |
| n                               | 749     | 749     | 749       | 749      |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

Table 4: The Joint Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of New Teaching Licenses

|                                          | (1)                 | (2)                | (3)                | (4)                  | (5)                 | (6)                 | (7)                 | (8)                  |
|------------------------------------------|---------------------|--------------------|--------------------|----------------------|---------------------|---------------------|---------------------|----------------------|
| Implement Evaluation                     | -2.505**<br>(0.861) |                    | -2.254*<br>(0.875) | -2.787***<br>(0.737) | -2.366**<br>(0.832) |                     | -1.754*<br>(0.832)  | -2.130**<br>(0.773)  |
| Eliminate Tenure                         |                     | -2.822+<br>(1.430) | -2.340<br>(1.540)  | -2.715+<br>(1.505)   |                     | -3.994**<br>(1.242) | -3.134*<br>(1.463)  | -2.297<br>(1.488)    |
| Implement Evaluation * Year_R (Eval)     |                     |                    |                    |                      | -1.341*<br>(0.571)  |                     | -1.495**<br>(0.515) | -1.724***<br>(0.485) |
| Year_R (Eval)                            |                     |                    |                    |                      | 0.330<br>(0.204)    |                     | 0.316<br>(0.206)    | 0.134<br>(0.219)     |
| Eliminate Tenure * Year_R (Tenure)       |                     |                    |                    |                      |                     | 1.217+<br>(0.614)   | 1.509*<br>(0.643)   | 1.538*<br>(0.660)    |
| Year_R (Tenure)                          |                     |                    |                    |                      |                     | -0.188<br>(0.384)   | -0.116<br>(0.318)   | -0.294<br>(0.297)    |
| % change relative to state mean (Eval)   | -17%                |                    | -15%               |                      |                     |                     |                     |                      |
| % change relative to state mean (Tenure) |                     | -19%               | -16%               |                      |                     |                     |                     |                      |
| n                                        | 749                 | 749                | 749                | 749                  | 749                 | 749                 | 749                 | 749                  |
| Standard Controls                        | Y                   | Y                  | Y                  | Y                    | Y                   | Y                   | Y                   | Y                    |
| Controls for Other Education Reforms     |                     |                    |                    | Y                    |                     |                     |                     | Y                    |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects. Controls for additional time-varying concurrent education reforms include the following: an indicator for collective bargaining coded as 1 if not mandatory and 0 if mandatory; an indicator for right-to-work coded as 1 if a state does not allow mandatory agency fees, 0 otherwise; an indicator for states that require new teachers to pass a basic skills test to obtain a licensure; an indicator for states that require new teachers to pass a content area test to obtain a licensure; an indicator for states that require new teachers to pass an exam testing professional knowledge to obtain a licensure; an indicator for states that adopted the Common Core State Standards; an indicator for states that administered Common Core aligned tests; and the proportion of teachers' total salary contributed to pension funds (i.e. employee contribution rate).

Table 5: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of Graduates from University-Based Bachelor's and Master's Teacher Preparation Programs

|                                         | (1)               | (2)               | (3)               | (4)               |
|-----------------------------------------|-------------------|-------------------|-------------------|-------------------|
| Panel A: High-Stakes Evaluation Reforms |                   |                   |                   |                   |
| Implement Evaluation                    | -0.518<br>(0.496) | -0.728<br>(0.488) | -0.416<br>(0.414) | -0.569<br>(0.418) |
| Implement Evaluation * Year_R           |                   |                   | -0.101<br>(0.246) | -0.092<br>(0.256) |
| Year_R                                  |                   |                   | -0.029<br>(0.131) | -0.083<br>(0.139) |
| % change relative to state mean         | -4%               | -5%               |                   |                   |
| Panel B: Tenure Reforms                 |                   |                   |                   |                   |
| Eliminate Tenure                        | -0.745<br>(0.659) | -1.116<br>(0.837) | -0.125<br>(0.664) | 0.143<br>(0.777)  |
| Eliminate Tenure * Year_R               |                   |                   | -0.049<br>(0.230) | -0.022<br>(0.230) |
| Year_R                                  |                   |                   | -0.080<br>(0.136) | -0.211<br>(0.141) |
| % change relative to state mean         | -6%               | -8%               |                   |                   |
| Controls                                |                   | Y                 |                   | Y                 |
| n                                       | 750               | 750               | 750               | 750               |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

Table 6: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of Graduates from University-Based Bachelor's and Master's Teacher Preparation Programs by Subject Area

|                                         | Non-Shortage Licensure Areas |                    |                   |                    |                    |                         |                   |                    | Shortage Licensure Areas |                   |                   |                    |                    |                    |                   |                   |
|-----------------------------------------|------------------------------|--------------------|-------------------|--------------------|--------------------|-------------------------|-------------------|--------------------|--------------------------|-------------------|-------------------|--------------------|--------------------|--------------------|-------------------|-------------------|
|                                         | Combined Non-Shortage Areas  |                    | Elementary        | English            | Social Studies     | Combined Shortage Areas |                   | Math               | Science                  | Special Education |                   |                    |                    |                    |                   |                   |
| Panel A: High-Stakes Evaluation Reforms |                              |                    |                   |                    |                    |                         |                   |                    |                          |                   |                   |                    |                    |                    |                   |                   |
| Implement Evaluation                    | -0.154<br>(0.195)            | -0.172<br>(0.185)  | -0.008<br>(0.185) | -0.108<br>(0.186)  | -0.124+<br>(0.066) | -0.068<br>(0.061)       | -0.019<br>(0.015) | 0.002<br>(0.014)   | -0.015<br>(0.100)        | 0.024<br>(0.103)  | 0.007<br>(0.019)  | 0.004<br>(0.019)   | -0.070*<br>(0.029) | -0.052+<br>(0.027) | 0.057<br>(0.075)  | 0.079<br>(0.075)  |
| Implement Evaluation * Year_R           |                              | 0.023<br>(0.122)   |                   | 0.119<br>(0.108)   |                    | -0.079*<br>(0.034)      |                   | -0.016<br>(0.011)  |                          | -0.060<br>(0.063) |                   | -0.027+<br>(0.014) |                    | -0.026+<br>(0.013) |                   | -0.005<br>(0.047) |
| Year_R                                  |                              | 0.006<br>(0.047)   |                   | 0.032<br>(0.043)   |                    | -0.014<br>(0.017)       |                   | -0.011*<br>(0.005) |                          | -0.009<br>(0.030) |                   | 0.011<br>(0.008)   |                    | -0.006<br>(0.005)  |                   | -0.014<br>(0.022) |
| % change relative to state mean         | -3%                          |                    | 0%                |                    | -23%               |                         | -6%               |                    | -1%                      |                   | 3%                |                    | -33%               |                    | 4%                |                   |
| Panel B: Tenure Reforms                 |                              |                    |                   |                    |                    |                         |                   |                    |                          |                   |                   |                    |                    |                    |                   |                   |
| Eliminate Tenure                        | -0.403<br>(0.251)            | -0.162<br>(0.241)  | -0.272<br>(0.219) | -0.176<br>(0.213)  | -0.127<br>(0.101)  | 0.034<br>(0.061)        | -0.004<br>(0.039) | -0.017<br>(0.026)  | -0.160<br>(0.195)        | 0.037<br>(0.127)  | -0.031<br>(0.028) | 0.015<br>(0.024)   | -0.032<br>(0.031)  | 0.021<br>(0.018)   | -0.121<br>(0.172) | 0.023<br>(0.113)  |
| Eliminate Tenure * Year_R               |                              | -0.230*<br>(0.094) |                   | -0.169+<br>(0.086) |                    | -0.042<br>(0.029)       |                   | -0.019<br>(0.012)  |                          | -0.069<br>(0.050) |                   | -0.012<br>(0.009)  |                    | 0.009<br>(0.010)   |                   | -0.057<br>(0.045) |
| Year_R                                  |                              | 0.032<br>(0.052)   |                   | 0.037<br>(0.052)   |                    | -0.014<br>(0.016)       |                   | 0.008+<br>(0.005)  |                          | -0.012<br>(0.033) |                   | -0.004<br>(0.009)  |                    | -0.012<br>(0.009)  |                   | -0.007<br>(0.026) |
| % change relative to state mean         | -9%                          |                    | -7%               |                    | -24%               |                         | -1%               |                    | -9%                      |                   | -14%              |                    | -15%               |                    | -9%               |                   |
| Controls                                | Y                            | Y                  | Y                 | Y                  | Y                  | Y                       | Y                 | Y                  | Y                        | Y                 | Y                 | Y                  | Y                  | Y                  | Y                 | Y                 |
| n                                       | 750                          | 750                | 750               | 750                | 740                | 740                     | 706               | 706                | 750                      | 750               | 720               | 720                | 702                | 702                | 750               | 750               |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*p<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

Table 7: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of Graduates from University-Based Bachelor's and Master's Teacher Preparation Programs by the Competitiveness of the Admissions Process

|                                         | Very Competitive College |         | Competitive College |         | Less Competitive College |         | Unranked |         |
|-----------------------------------------|--------------------------|---------|---------------------|---------|--------------------------|---------|----------|---------|
| Panel A: High-Stakes Evaluation Reforms |                          |         |                     |         |                          |         |          |         |
| Implement Evaluation                    | -0.115                   | -0.059  | -0.408              | -0.348  | 0.221                    | 0.179   | -0.418+  | -0.374  |
|                                         | (0.145)                  | (0.110) | (0.442)             | (0.373) | (0.133)                  | (0.121) | (0.223)  | (0.224) |
| Implement Evaluation * Year_R           |                          | -0.062  |                     | -0.106  |                          | 0.184*  |          | -0.199  |
|                                         |                          | (0.090) |                     | (0.243) |                          | (0.078) |          | (0.140) |
| Year_R                                  |                          | -0.019  |                     | -0.008  |                          | -0.030  |          | 0.033   |
|                                         |                          | (0.034) |                     | (0.127) |                          | (0.025) |          | (0.061) |
| % change relative to state mean         | -4%                      |         | -6%                 |         | 8%                       |         | -41%     |         |
| Panel B: Tenure Reforms                 |                          |         |                     |         |                          |         |          |         |
| Eliminate Tenure                        | -0.087                   | -0.203  | -0.303              | 0.160   | -0.317                   | -0.345  | -0.262   | 0.473   |
|                                         | (0.254)                  | (0.188) | (0.611)             | (0.525) | (0.197)                  | (0.232) | (0.314)  | (0.444) |
| Eliminate Tenure * Year_R               |                          | 0.106*  |                     | 0.035   |                          | -0.005  |          | -0.041  |
|                                         |                          | (0.050) |                     | (0.217) |                          | (0.083) |          | (0.087) |
| Year_R                                  |                          | -0.014  |                     | -0.091  |                          | 0.006   |          | -0.114  |
|                                         |                          | (0.042) |                     | (0.097) |                          | (0.040) |          | (0.075) |
| % change relative to state mean         | -3%                      |         | -5%                 |         | -12%                     |         | -26%     |         |
| Controls                                | Y                        | Y       | Y                   | Y       | Y                        | Y       | Y        | Y       |
| n                                       | 750                      | 750     | 750                 | 750     | 750                      | 750     | 750      | 750     |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

Table 8: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Average Freshman SAT Scores at the Universities Where Candidates Received their Bachelor's or Master's Teaching Degrees

|                               | SAT Math                                |          |                 |          |
|-------------------------------|-----------------------------------------|----------|-----------------|----------|
|                               | 25th Percentile                         |          | 75th Percentile |          |
|                               | Panel A: High-Stakes Evaluation Reforms |          |                 |          |
| Implement Evaluation          | -2.750                                  | -1.910   | -1.496          | -1.588   |
|                               | (2.005)                                 | (1.550)  | (1.424)         | (1.165)  |
| Implement Evaluation * Year_R |                                         | -0.212   |                 | 0.869    |
|                               |                                         | (1.125)  |                 | (0.881)  |
| Year_R                        |                                         | -0.775   |                 | -0.364   |
|                               |                                         | (0.822)  |                 | (0.676)  |
|                               | Panel B: Tenure Reforms                 |          |                 |          |
| Eliminate Tenure              | 5.771+                                  | 0.697    | 2.026           | 0.674    |
|                               | (2.959)                                 | (1.811)  | (2.981)         | (1.779)  |
| Eliminate Tenure * Year_R     |                                         | 4.666*** |                 | 3.153*** |
|                               |                                         | (1.127)  |                 | (0.845)  |
| Year_R                        |                                         | -0.092   |                 | -0.500   |
|                               |                                         | (0.443)  |                 | (0.386)  |
| Controls                      | Y                                       | Y        | Y               | Y        |
| n                             | 600                                     | 600      | 600             | 600      |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

Table 9: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of Graduates from University-Based Bachelor's and Master's Teacher Preparation Programs by the Competitiveness of the Admissions Process

|                                         | Female            |                   | Male              |                   | Asian             |                    | Black              |                   | Hispanic          |                    | White             |                   |
|-----------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|--------------------|-------------------|-------------------|--------------------|-------------------|-------------------|
| Panel A: High-Stakes Evaluation Reforms |                   |                   |                   |                   |                   |                    |                    |                   |                   |                    |                   |                   |
| Implement Evaluation                    | -0.593<br>(0.414) | -0.448<br>(0.351) | -0.136<br>(0.087) | -0.121<br>(0.080) | -0.024<br>(0.029) | -0.007<br>(0.013)  | -0.071<br>(0.085)  | -0.008<br>(0.087) | -0.037<br>(0.030) | -0.030<br>(0.025)  | -0.488<br>(0.350) | -0.378<br>(0.328) |
| Implement Evaluation * Year_R           |                   | -0.046<br>(0.211) |                   | -0.047<br>(0.048) |                   | 0.000<br>(0.006)   |                    | 0.016<br>(0.041)  |                   | 0.009<br>(0.019)   |                   | -0.067<br>(0.174) |
| Year_R                                  |                   | -0.088<br>(0.119) |                   | 0.005<br>(0.023)  |                   | -0.012<br>(0.013)  |                    | -0.050<br>(0.036) |                   | -0.008<br>(0.010)  |                   | -0.056<br>(0.075) |
| % change relative to state mean         | -6%               |                   | -5%               |                   | -10%              |                    | -9%                |                   | -7%               |                    | -5%               |                   |
| Panel B: Tenure Reforms                 |                   |                   |                   |                   |                   |                    |                    |                   |                   |                    |                   |                   |
| Eliminate Tenure                        | -0.936<br>(0.709) | 0.169<br>(0.655)  | -0.180<br>(0.137) | -0.026<br>(0.131) | 0.000<br>(0.030)  | -0.003<br>(0.013)  | -0.331+<br>(0.185) | -0.224<br>(0.282) | -0.047<br>(0.031) | 0.048+<br>(0.027)  | -0.924<br>(0.809) | -0.727<br>(0.939) |
| Eliminate Tenure * Year_R               |                   | -0.064<br>(0.198) |                   | 0.042<br>(0.042)  |                   | -0.009*<br>(0.004) |                    | -0.051<br>(0.062) |                   | -0.026*<br>(0.011) |                   | -0.175<br>(0.185) |
| Year_R                                  |                   | -0.171<br>(0.118) |                   | -0.040<br>(0.028) |                   | 0.003<br>(0.005)   |                    | -0.002<br>(0.041) |                   | -0.008<br>(0.006)  |                   | 0.022<br>(0.121)  |
| % change relative to state mean         | -9%               |                   | -6%               |                   | 0%                |                    | -44%               |                   | -9%               |                    | -9%               |                   |
| Controls                                | Y                 | Y                 | Y                 | Y                 | Y                 | Y                  | Y                  | Y                 | Y                 | Y                  | Y                 | Y                 |
| n                                       | 750               | 750               | 750               | 750               | 750               | 750                | 750                | 750               | 750               | 750                | 750               | 750               |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table 2. All models include state and region-by-year fixed effects.

## Figures

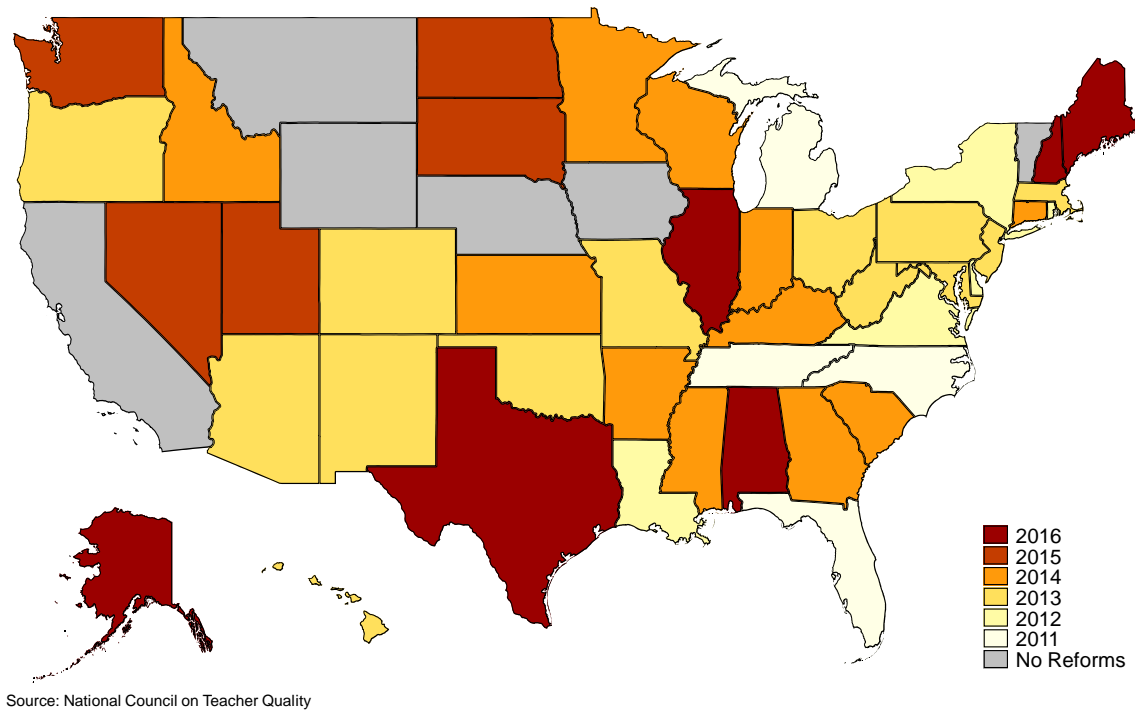


Figure 1: The timing of statewide full implementation of new high-stakes teacher evaluation systems.

Notes: Years represent the fall academic year in which the new systems were first fully implemented statewide.



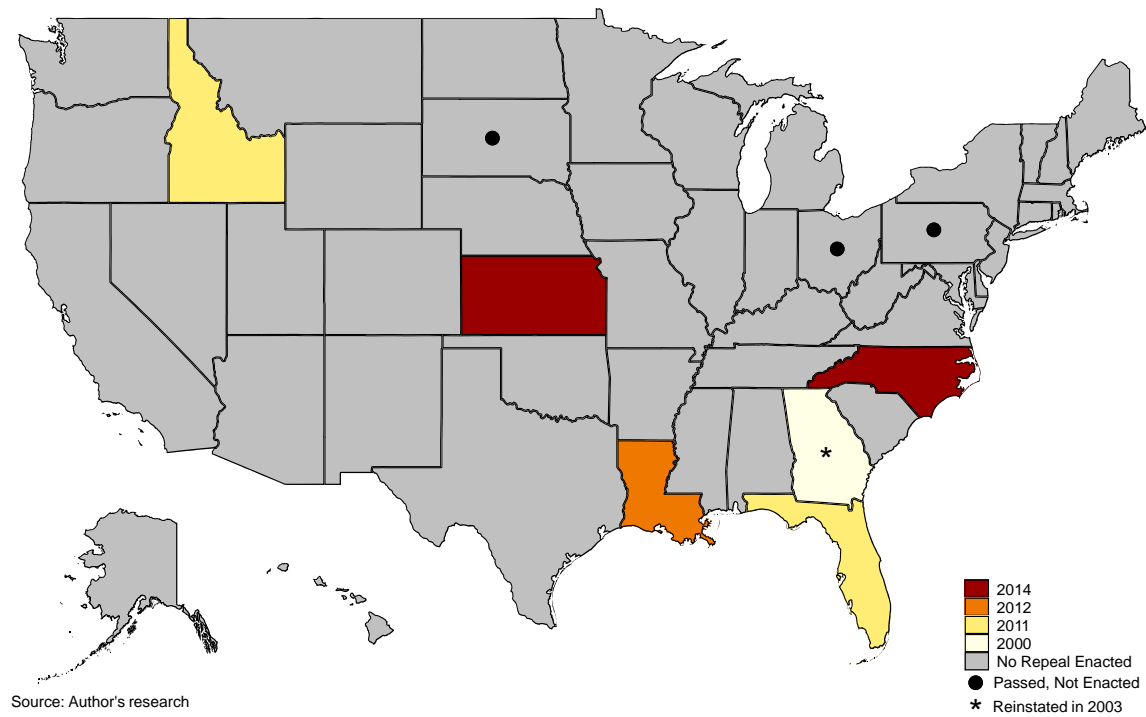


Figure 2: The timing of state legislative repeal or effective elimination of teacher tenure for new teachers.

Notes: Years represent the calendar year in which a law was passed.

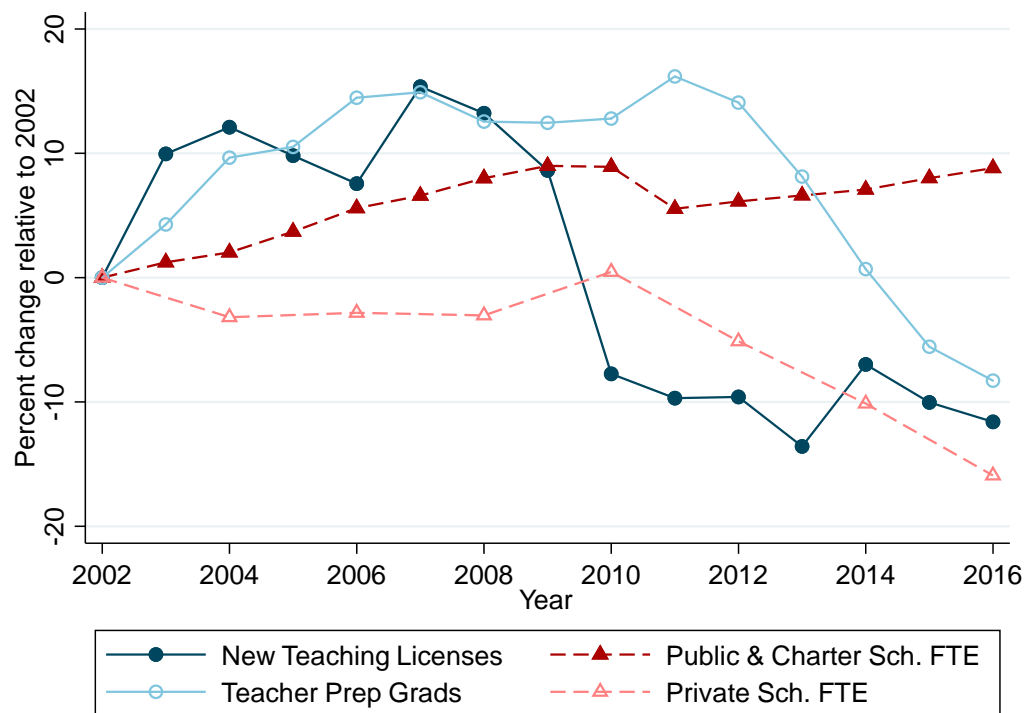


Figure 3: National Trends in the Teacher Labor Market.

Notes: Data on the number of teaching licenses issued come from Title II data provided by states to the U.S. DOE. A reporting year for Title II is from September 1<sup>st</sup> to August 31<sup>st</sup> for the latter year. Data on the number of traditional teacher preparation program graduates at post-secondary institutions is from the IPEDS Survey administered by NCES. IPEDS reports graduates in a 12 month period up to the spring and summer of the given year. Data on total K-12 public school, charter school, and private school teachers are based on total full-time equivalent (FTE) teaching positions from the NCES common core state-level files using spring academic year. Public and charter school FTE is a lower bound estimate as charter school FTE data are missing in 6 percent of state-year cells. All data are scaled by their 2002 values such that trends represent the percent change in a given measure relative to 2002 levels. The baseline 2002 values are 277,696 for licenses, 211,764 for teacher preparation program graduates, 2,994,575 for public school teachers, 22,785 for charter school teachers, and 387,495 for private school teachers.

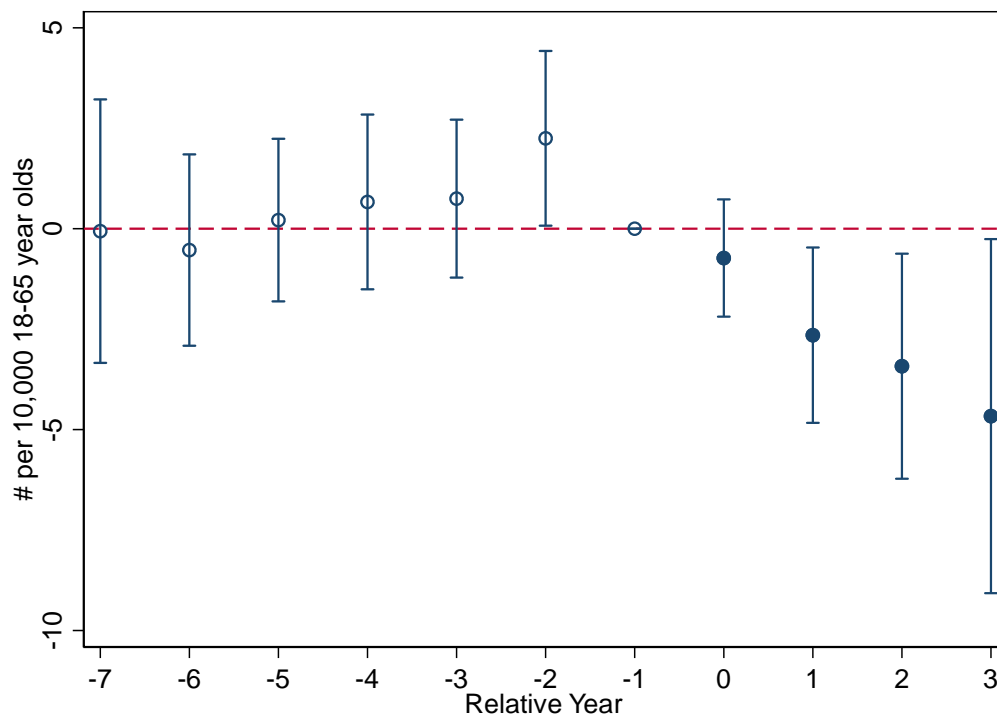


Figure 4: Event study depicting effect of evaluation reforms on the number of new teaching licensures

Notes: Point estimates for years pre- (hollow dot) and post-reform (solid dot) and corresponding 95 percent confident intervals are derived from an event study model that includes time-varying state-level controls: real average hourly wages in the private sector, unemployment rates, real state tax revenue per capita, the fraction of the population ages 5 to 17 that are White, Black, Hispanic, and living below the federal poverty line, and measures specific to K-12 public education: real average teacher salaries and real current per pupil expenditures, and region-by-year fixed effects. Estimates depicted in this figure are reported in Appendix Table A5.

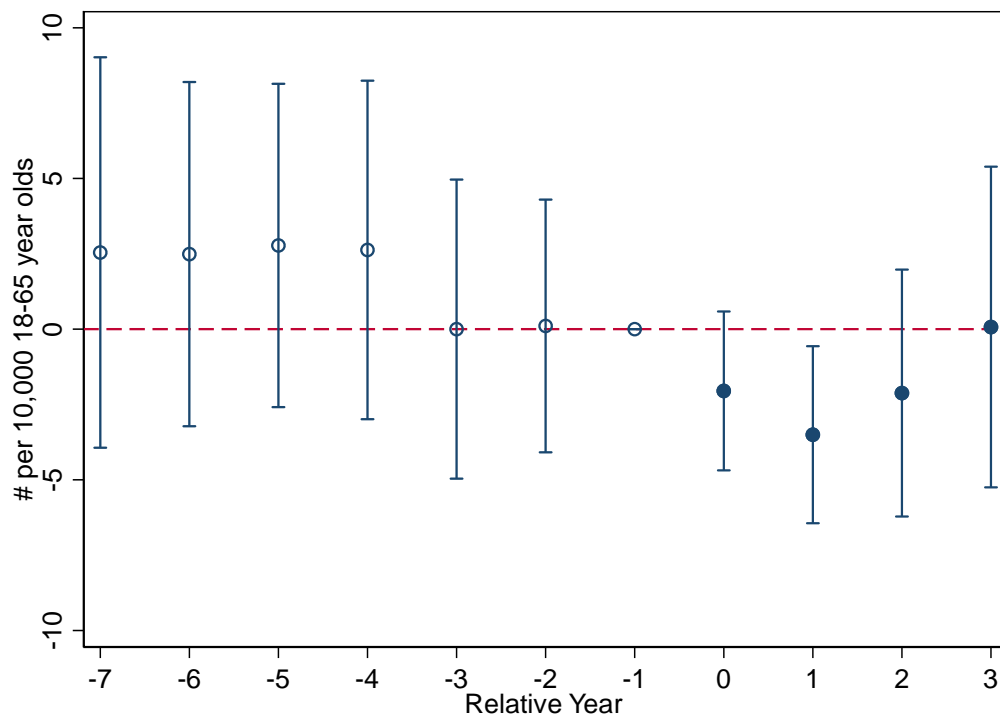


Figure 5. Event study depicting effect of repealing tenure on the number of new teaching licensures

Notes: See Figure 4 for details.

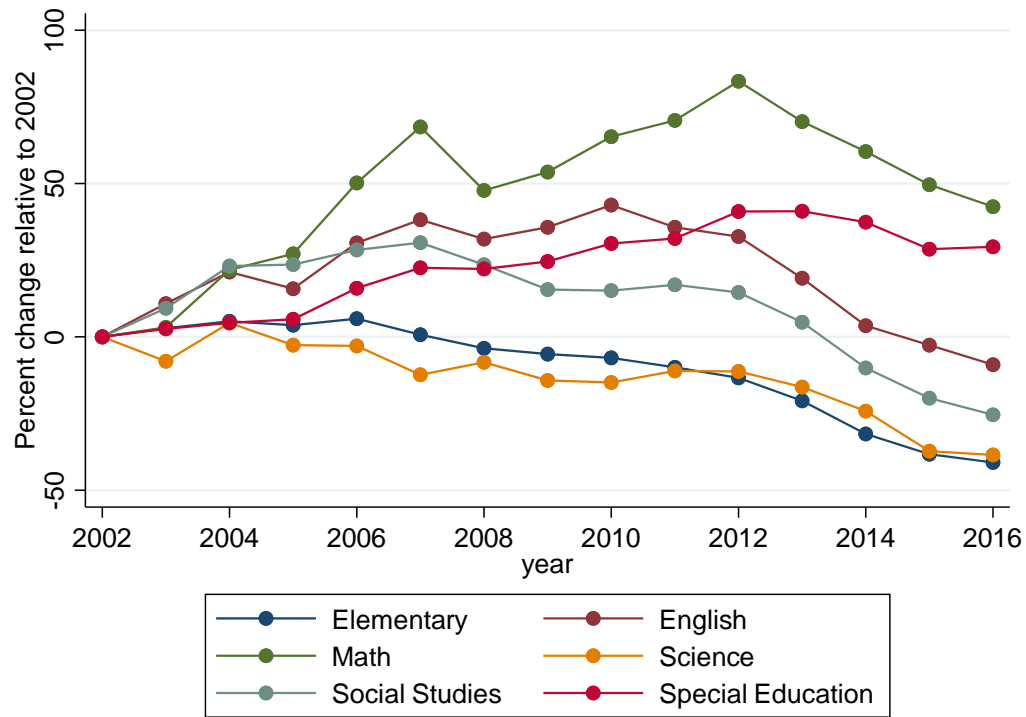


Figure 6: National trends in the number of university-based degree completers by subject area

Notes: Data on the number of traditional teacher preparation program graduates at post-secondary institutions that receive financial aid is from the IPEDS Survey administered by NCES. All data are scaled by their 2002 values such that trends represent the percent change in a given measure relative to 2002 levels.

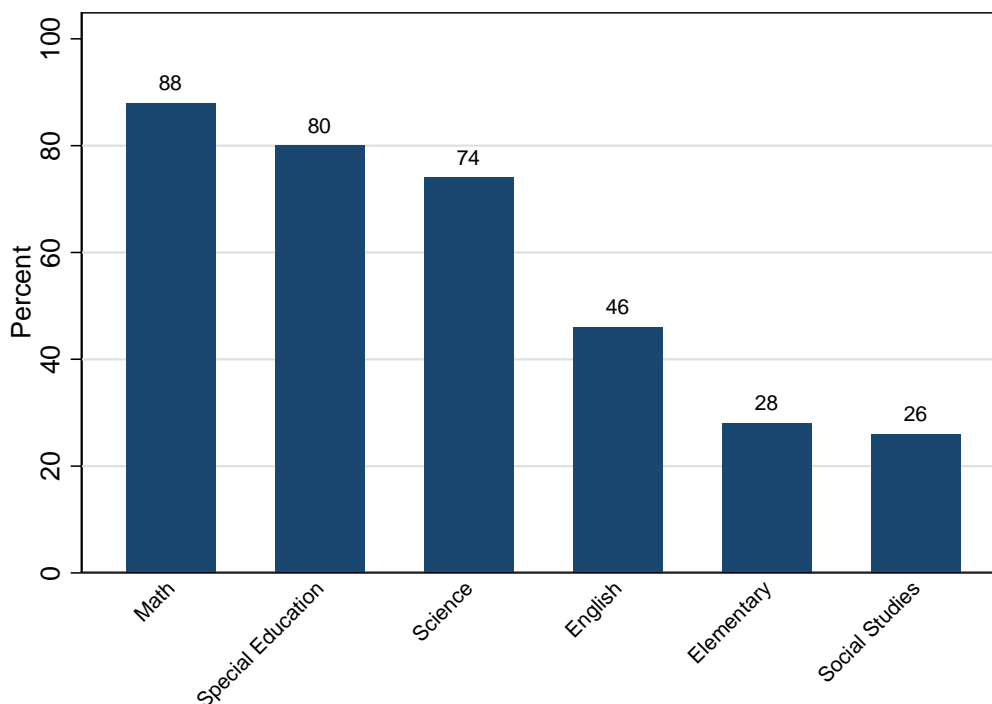


Figure 7: State Teacher Shortage Areas by Licensure Type in 2017-18 as Designated by the U.S. Secretary of Education.

Notes: A teacher shortage area (TSA) is “an area of specific grade, subject matter or discipline classification, or a geographic area in which the Secretary [of Education] determines that there is an inadequate supply of elementary or secondary school teachers” (34 CFR 682.210(q)(8)(vii)). States submit a TSA proposal for consideration and are evaluated based on a few criteria, including a requirement that the TSAs may not account for more than five percent of full time teaching positions in the state. Teachers who teach in federally recognized TSAs make themselves eligible for deferment or even cancellation of their federal loans, including Perkins and Stafford Loans. Detailed notes on coding equivalent licensure areas across states available upon request.

## Appendix A

### Appendix Tables

Table A1: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms from Event Study Models

| Outcome   | Teacher Licensures Issued |                    |                    |                    | Teacher Prep Program Completers |                    |                   |                    |
|-----------|---------------------------|--------------------|--------------------|--------------------|---------------------------------|--------------------|-------------------|--------------------|
| Predictor | Evaluation                |                    | Tenure             |                    | Evaluation                      |                    | Tenure            |                    |
| tm7+      | 0.516<br>(1.754)          | -0.061<br>(1.672)  | 1.833<br>(3.228)   | 2.544<br>(3.305)   | 0.545<br>(1.091)                | 1.003<br>(1.078)   | 1.178<br>(0.988)  | 2.387*<br>(1.131)  |
| tm6       | 0.056<br>(1.344)          | -0.533<br>(1.215)  | 3.595<br>(2.484)   | 2.489<br>(2.914)   | 0.689<br>(0.919)                | 1.092<br>(0.889)   | 0.963<br>(0.772)  | 2.104*<br>(0.990)  |
| tm5       | 0.540<br>(1.220)          | 0.213<br>(1.032)   | 4.113<br>(2.509)   | 2.776<br>(2.737)   | 0.506<br>(0.860)                | 0.824<br>(0.840)   | 1.028<br>(0.700)  | 1.987*<br>(0.904)  |
| tm4       | 0.955<br>(1.262)          | 0.665<br>(1.110)   | 3.758<br>(2.793)   | 2.627<br>(2.866)   | 0.520<br>(0.700)                | 0.739<br>(0.691)   | 1.255+<br>(0.686) | 2.049*<br>(0.835)  |
| tm3       | 0.845<br>(1.103)          | 0.748<br>(1.003)   | 0.684<br>(2.476)   | 0.000<br>(2.532)   | 0.303<br>(0.450)                | 0.430<br>(0.467)   | 1.137*<br>(0.522) | 1.742**<br>(0.649) |
| tm2       | 2.281+<br>(1.136)         | 2.249*<br>(1.109)  | 0.397<br>(2.230)   | 0.106<br>(2.139)   | 0.267<br>(0.233)                | 0.361<br>(0.241)   | 0.896+<br>(0.481) | 1.324*<br>(0.586)  |
| t0        | -0.802<br>(0.703)         | -0.730<br>(0.745)  | -2.115+<br>(1.252) | -2.050<br>(1.345)  | -0.363<br>(0.274)               | -0.476+<br>(0.281) | 0.361<br>(0.430)  | 0.621<br>(0.590)   |
| t1        | -2.679*<br>(1.070)        | -2.650*<br>(1.113) | -2.874+<br>(1.465) | -3.503*<br>(1.498) | -0.406<br>(0.521)               | -0.684<br>(0.547)  | 0.370<br>(0.717)  | 0.509<br>(0.925)   |
| t2        | -3.558*<br>(1.370)        | -3.423*<br>(1.429) | -1.323<br>(2.050)  | -2.121<br>(2.089)  | -0.306<br>(0.801)               | -0.628<br>(0.831)  | -0.232<br>(0.739) | -0.020<br>(1.022)  |
| t3+       | -4.828*<br>(2.095)        | -4.665*<br>(2.248) | 0.626<br>(2.895)   | 0.070<br>(2.715)   | -0.691<br>(1.093)               | -0.971<br>(1.134)  | -0.121<br>(0.684) | -0.111<br>(1.004)  |
| Controls  | Y                         |                    | Y                  |                    | Y                               |                    | Y                 |                    |
| n         | 749                       | 749                | 749                | 749                | 750                             | 750                | 750               | 750                |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*p<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls include real average hourly wages in the private sector, unemployment rates, real state tax revenue per capita, the fraction of the population ages 5 to 17 that are White, Black, Hispanic, and living below the federal poverty line, and measures specific to the K-12 public education: real average teacher salaries and real current per pupil expenditures. All models include state and region-by-year fixed effects.

Table A2: The Effect of Eliminating Tenure on the Number of New Teaching Licenses in Models that Drop One State

|                           | Without | Florida | Without<br>Georgia | Without<br>Idaho | Without<br>Kansas | Without<br>Louisiana | Without North Carolina |          |         |           |         |          |
|---------------------------|---------|---------|--------------------|------------------|-------------------|----------------------|------------------------|----------|---------|-----------|---------|----------|
| Eliminate Tenure          | -4.064+ | -3.087* | -1.853             | -3.092*          | -2.802+           | -4.032**             | -3.134*                | -3.882** | -3.572* | -4.773*** | -1.772  | -4.655** |
|                           | (2.319) | (1.472) | (1.782)            | (1.166)          | (1.487)           | (1.333)              | (1.462)                | (1.363)  | (1.497) | (1.238)   | (1.208) | (1.592)  |
| Eliminate Tenure * Year_R |         | 0.153   |                    | 1.617**          |                   | 1.241+               |                        | 1.519*   |         | 1.328*    |         | 0.627    |
|                           |         | (0.828) |                    | (0.503)          |                   | (0.635)              |                        | (0.589)  |         | (0.622)   |         | (0.532)  |
| Year_R                    |         | -0.201  |                    | -0.243           |                   | -0.192               |                        | -0.368   |         | -0.234    |         | 0.258    |
|                           |         | (0.393) |                    | (0.372)          |                   | (0.412)              |                        | (0.397)  |         | (0.445)   |         | (0.279)  |
| Controls                  | Y       | Y       | Y                  | Y                | Y                 | Y                    | Y                      | Y        | Y       | Y         | Y       | Y        |
| N                         | 734     | 734     | 734                | 734              | 734               | 734                  | 734                    | 734      | 734     | 734       | 734     | 734      |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.



Table A3: The Effect of High-Stakes Teacher Evaluation Reforms on the Number of New Teaching Licenses in a Restricted Estimation Sample that Excludes States that Repealed Tenure

|                                 | Full                | Restricted         | Full                | Restricted         |
|---------------------------------|---------------------|--------------------|---------------------|--------------------|
| Implement Evaluation            | -2.505**<br>(0.861) | -2.039*<br>(0.958) | -2.366**<br>(0.832) | -1.819+<br>(0.935) |
| Implement Evaluation * Year_R   |                     |                    | -1.341*<br>(0.571)  | -1.378*<br>(0.536) |
| Year_R                          |                     |                    | 0.330<br>(0.204)    | 0.329<br>(0.216)   |
| % change relative to state mean | -17%                | -14%               |                     |                    |
| n                               | 749                 | 659                | 749                 | 659                |
| Controls                        | Y                   | Y                  | Y                   | Y                  |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

Table A4: Falsification Tests with Certified Public Accountant Exam Takers

| Panel A: High-Stakes Evaluation Reforms |         |         |         |         |
|-----------------------------------------|---------|---------|---------|---------|
| Implement Evaluation                    | 0.011   | 0.001   | -0.091  | -0.076  |
|                                         | (0.108) | (0.109) | (0.099) | (0.104) |
| Implement Evaluation * Year_R           |         |         | 0.057   | 0.053   |
|                                         |         |         | (0.079) | (0.095) |
| Year_R                                  |         |         | 0.069   | 0.062   |
|                                         |         |         | (0.093) | (0.058) |
| % change relative to state mean         | 0%      | 0%      |         |         |
| Panel B: Tenure Reforms                 |         |         |         |         |
| Eliminate Tenure                        | 0.155   | 0.159   | 0.012   | 0.041   |
|                                         | (0.101) | (0.245) | (0.080) | (0.153) |
| Eliminate Tenure * Year_R               |         |         | 0.038   | 0.071   |
|                                         |         |         | (0.048) | (0.078) |
| Year_R                                  |         |         | 0.015   | 0.003   |
|                                         |         |         | (0.032) | (0.061) |
| % change relative to state mean         | 6%      | 6%      |         |         |
| Controls                                |         | Y       |         | Y       |
| n                                       | 550     | 550     | 550     | 550     |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

Table A5: Test of Treatment Spillover in Contiguous States due to Endogenous Mobility

|                                    | Restricted panel from<br>2002 to 2012 |                    | Full Panel         |                    |
|------------------------------------|---------------------------------------|--------------------|--------------------|--------------------|
| Implement Evaluation (Eval)        | -4.418**<br>(1.543)                   | -3.179*<br>(1.570) |                    |                    |
| Eval Bordering Comparison States   |                                       | 1.465<br>(1.352)   |                    |                    |
| Eliminate Tenure                   |                                       |                    | -2.822+<br>(1.430) | -3.069*<br>(1.313) |
| Tenure Bordering Comparison States |                                       |                    |                    | -0.375<br>(0.911)  |
| % change relative to state mean    | -30%                                  | -21%               | -19%               | -21%               |
| Controls                           | Y                                     | Y                  | Y                  | Y                  |
| n                                  | 550                                   | 550                | 749                | 749                |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

Table A6: Alternative Specifications for the Effect of High-Stakes Teacher Evaluation and Tenure Reforms on the Number of New Teaching Licenses

| Outcome specification                   | Scaled per<br>number of<br>18-65 year<br>olds per<br>10,000 | Scaled per<br>number of<br>22-25 year<br>olds per<br>1,000 | Scaled per<br>number of<br>18-65 year<br>olds per<br>10,000 | Logged,<br>controlling<br>for logged<br>number of<br>18-65 year<br>olds per<br>10,000 | Scaled per<br>number of<br>18-65 year<br>olds per<br>10,000 | Scaled per<br>number of<br>22-25 year<br>olds per<br>1,000 | Scaled per<br>number of<br>18-65 year<br>olds per<br>10,000 | Logged,<br>controlling<br>for logged<br>number of<br>18-65 year<br>olds per<br>10,000 |
|-----------------------------------------|-------------------------------------------------------------|------------------------------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------------------------------|------------------------------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------------------------------|
|                                         | Number of<br>18-65 year<br>olds per<br>10,000               | Number of<br>22-25 year<br>olds per<br>1,000               | No weights                                                  | No weights                                                                            | Number of<br>18-65 year<br>olds per<br>10,000               | Number of<br>22-25 year<br>olds per<br>1,000               | No weights                                                  | No weights                                                                            |
| Weights                                 | (1)                                                         | (2)                                                        | (3)                                                         | (4)                                                                                   | (5)                                                         | (6)                                                        | (7)                                                         | (8)                                                                                   |
| Panel A: High-Stakes Evaluation Reforms |                                                             |                                                            |                                                             |                                                                                       |                                                             |                                                            |                                                             |                                                                                       |
| Implement Evaluation                    | -2.505**<br>(0.861)                                         | -3.142**<br>(1.021)                                        | -2.190*<br>(1.031)                                          | -0.145+<br>(0.072)                                                                    | -2.366**<br>(0.832)                                         | -3.018**<br>(0.993)                                        | -2.371*<br>(1.009)                                          | -0.153*<br>(0.071)                                                                    |
| Implement Evaluation * Year_R           |                                                             |                                                            |                                                             |                                                                                       | -1.341*<br>(0.571)                                          | -1.622*<br>(0.660)                                         | -0.899+<br>(0.526)                                          | -0.073+<br>(0.041)                                                                    |
| Year_R                                  |                                                             |                                                            |                                                             |                                                                                       | 0.330<br>(0.204)                                            | 0.423+<br>(0.230)                                          | 0.308<br>(0.210)                                            | 0.021<br>(0.014)                                                                      |
| % change relative to state mean         | -17%                                                        | -18%                                                       | -16%                                                        |                                                                                       |                                                             |                                                            |                                                             |                                                                                       |
| Panel B: Tenure Reforms                 |                                                             |                                                            |                                                             |                                                                                       |                                                             |                                                            |                                                             |                                                                                       |
| Eliminate Tenure                        | -2.822+<br>(1.430)                                          | -3.559*<br>(1.690)                                         | -2.367+<br>(1.265)                                          | -0.204*<br>(0.096)                                                                    | -3.994**<br>(1.242)                                         | -4.657**<br>(1.519)                                        | -3.248+<br>(1.692)                                          | -0.242+<br>(0.127)                                                                    |
| Eliminate Tenure * Year_R               |                                                             |                                                            |                                                             |                                                                                       | 1.217+<br>(0.614)                                           | 1.466*<br>(0.721)                                          | 0.671<br>(0.676)                                            | 0.065<br>(0.053)                                                                      |
| Year_R                                  |                                                             |                                                            |                                                             |                                                                                       | -0.188<br>(0.384)                                           | -0.280<br>(0.451)                                          | -0.044<br>(0.301)                                           | -0.012<br>(0.022)                                                                     |
| % change relative to state mean         | -19%                                                        | -21%                                                       | -18%                                                        |                                                                                       |                                                             |                                                            |                                                             |                                                                                       |
| Controls                                | Y                                                           | Y                                                          | Y                                                           | Y                                                                                     | Y                                                           | Y                                                          | Y                                                           | Y                                                                                     |
| n                                       | 749                                                         | 749                                                        | 749                                                         | 749                                                                                   | 749                                                         | 749                                                        | 749                                                         | 749                                                                                   |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

Table A7: The Effect of High-Stakes Teacher Evaluation and Tenure Reforms on Public Teacher Salaries

|                                         | NCES Average Salary |                      | ACS Adjusted Average Salary |                     |
|-----------------------------------------|---------------------|----------------------|-----------------------------|---------------------|
| Panel A: High-Stakes Evaluation Reforms |                     |                      |                             |                     |
| Implement Evaluation                    | -205.39<br>(558.22) | 100.59<br>(498.43)   | -446.62+<br>(232.71)        | -254.28<br>(185.22) |
| Implement Evaluation * Year_R           |                     | 61.94<br>(264.41)    |                             | -273.36<br>(187.56) |
| Year_R                                  |                     | -239.24+<br>(127.56) |                             | -83.30<br>(94.41)   |
| % change relative to state mean         | 0%                  |                      | -1%                         |                     |
| Panel B: Tenure Reforms                 |                     |                      |                             |                     |
| Eliminate Tenure                        | 74.76<br>(793.81)   | -474.73<br>(610.62)  | -118.48<br>(351.05)         | 578.04<br>(363.18)  |
| Eliminate Tenure* Year_R                |                     | 625.47+<br>(354.12)  |                             | 91.34<br>(212.28)   |
| Year_R                                  |                     | -101.62<br>(224.49)  |                             | -165.88<br>(130.67) |
| % change relative to state mean         | 0%                  |                      | 0%                          |                     |
| Controls                                | Y                   | Y                    | Y                           | Y                   |
| n                                       | 750                 | 750                  | 600                         | 600                 |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

## Appendix B

Table B1: State Legislation and Court Cases Restricting or Repealing Teacher Tenure, 2002-2016

| State   | Summary of Changes                                                                                                                                                                                                                                                                           | Bill(s) / Case                                    | Date Passed / Filed                          | Date Effective for Tenure Elimination | Went into Effect? | Date Overturned | How was it Overturned?                |
|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|----------------------------------------------|---------------------------------------|-------------------|-----------------|---------------------------------------|
| Florida | Governor Scott signed two important tenure-related laws: SB 736 required annual contracts for new teachers for the duration of their careers and salaries based on evaluations. HB 7087 prevented any teacher who did not already have tenure from receiving it.                             | Senate Bill 736 and House Bill 7087               | SB 736: March 24, 2011; HB 7087: May 5, 2011 | July 1, 2011                          | Yes               | N/A             | N/A                                   |
| Georgia | Governor Barnes signed and supported reform through House Bill 1187 (in which the opportunity to obtain tenure was eliminated for teachers hired after July 1, 2000). After little improvement, Governor Barnes was ousted and replaced with Governor Purdue, who promptly repealed the law. | House Bill 1187 (A+ Education Reform Act of 2000) | April 25, 2000                               | July 1, 2000                          | Yes               | June 4, 2003    | Legislative Repeal in Senate Bill 193 |

|           |                                                                                                                                                                                                                                                                                                              |                               |                |              |     |                  |                                                      |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------|----------------|--------------|-----|------------------|------------------------------------------------------|
| Idaho     | SB 1108, 1110 and 1194, or the "Luna Laws," eliminated tenure for new teachers, reworked teacher pay around performance and required supplying computers to every high school student. Opposition to these laws was swift after the laws went into effect; voters repealed the laws by a margin of over 50%. | Senate Bills 1108, 1110, 1194 | March 1, 2011  | July 1, 2011 | Yes | November 6, 2012 | Propositions 1, 2, & 3                               |
| Kansas    | In 2014, Governor Brownback signed HB 2506 redefining the word "teacher" so that due process procedures for public employees did not apply. Individual districts differ on the implementation of HB 2506.                                                                                                    | House Bill 2506               | July 1, 2014   | July 1, 2014 | Yes | N/A              | N/A                                                  |
| Louisiana | Act 1, passed in April 2012, increased the requirements for teacher tenure. To receive tenure, teachers must receive a highly-effective Compass rating five out of six consecutive years, an exceedingly difficult task. Tenure is immediately revoked if a teacher ever receives an ineffective rating.     | House Bill 974 (Act 1)        | April 18, 2012 | July 1, 2012 | Yes | N/A              | N/A (Upheld by Louisiana Supreme Court in Oct. 2014) |

|                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                 |                |               |     |                  |                                                                                                                                                                                                                                                                                                                |
|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----------------|---------------|-----|------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| North Carolina | Legislators in North Carolina revoked teacher tenure for those who already had it and for new teachers, in the annual 2013 budgeting process. The act of revoking tenure for those who already had career status (which would have been phased out completely by June 30, 2018) was overturned by the North Carolina Supreme Court in April 2016, but the new teacher provision was upheld, although individual districts can differ.                                                                                                                   | Senate Bill 402 | July 26, 2013  | July 1, 2014  | Yes | April 15, 2016   | Partially overturned by North Carolina Court of Appeals June 2, 2015, and then later ruled unconstitutional by the NC Supreme Court in 2016, which ruled the state cannot remove tenure from those who earned it before the law was passed, but can withhold this protection from those hired since July 2013. |
| Ohio           | Senate Bill 5 (2011), which eliminated awarding future "continuing contracts" (tenure) to teachers, was signed into law March 30, 2011 (with an effective date of June 30, 2011), but was overturned by Referendum Issue 2 on November 8, 2011. The law did not go into effect due to Ohio Constitution Article II §1c, which stipulates that "no such law [being petitioned and 'signed by six per centum of the electors of the state'], section or item shall go into effect until and unless approved by a majority of those voting upon the same." | Senate Bill 5   | March 31, 2011 | July 31, 2011 | No  | November 8, 2011 | Voter Referendum, Issue 2                                                                                                                                                                                                                                                                                      |



|              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                 |                   |               |    |                  |                                                |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|-------------------|---------------|----|------------------|------------------------------------------------|
| Pennsylvania | House Bill 805, first drafted in May 2015 and passed in May 2016 with a vote of 95-94, was vetoed by Governor Tom Wolf on May 18, 2016. The bill amended the Public School Code of 1949, changing the contracts of school teachers in aspects of receiving tenure, collective bargaining agreements, and timing of appeal and suspension. Governor Wolf cited that, while it is in everyone's best interest to improve accountability in education, HB 805 "relies heavily on a single score from the teacher evaluation system, as opposed to using the entire method of evaluation." | House Bill 805  | May 17, 2016      | June 30, 2016 | No | May 18, 2016     | Vetoed by Governor Tom Wolf                    |
| South Dakota | South Dakota passed HB 1234 in February 2012, which eliminated state mandates for tenure for teachers who had not achieved tenure by July 1, 2016, although individual districts would be able to continue offering tenure. The law was repealed by referendum the same year, which means that the tenure change never went into effect.                                                                                                                                                                                                                                               | House Bill 1234 | February 29, 2012 | July 1, 2016  | No | November 6, 2012 | Referred Law 16: Referendum on November ballot |

## Appendix C

### Identifying Graduates of Teacher Preparation Programs and by Subject Areas using CIPS Codes

The Integrated Post-Secondary Education Data System (IPEDS) includes the results of eleven surveys conducted annually by the U.S. Department of Education's National Center for Education Statistics (NCES). All institutions must submit race and ethnicity data if they receive, are applicants for, or expect to be applicants for federal financial assistance as defined under the DOE's regulations implementing Title VI of the Civil Rights Act of 1964, Title IX of the Education Amendments of 1972, and the Carl D. Perkins Education Act. These surveys thus capture the substantive universe of post-secondary schools, colleges, and universities.

In this paper, we use the entire "Completions" survey available for download on the IPEDS website. Each observation in the file corresponds to the completers of a particular academic program, identified by classification of instruction program (CIP) code, and the award level (e.g. bachelor's degree, master's degree, etc.) of that particular program for each reporting institution.

The NCES has devised a six-digit taxonomy for organizing academic programs called CIP codes. The first two digits of the code correspond to a broad area of study. For instance, all majors under the two-digit CIP category "13" are majors within the education field. These majors include not only education programs designed to prepare individuals to be teachers, but also teaching assistants preparation programs (13.1501) and programs for education program evaluators (e.g. Education Evaluation and Research, 13.0601), among others. The next two digits in the CIP code cluster majors with similar instructional content within the broad area of study. All majors with a 13.13 CIP code sequence, for example, are "teacher education or professional development majors within specific subject areas" within the broad field of education. The final two digits are unique to each specific major that falls within the specific subject area. For example, a "structural engineering" major has a 14.0803 CIP code ("14" corresponds to engineering, "08" correspond to civil engineering, and the final "03" is unique to structural engineering). NCES adds and removes CIP codes regularly. We created consistent categories for all six-digit codes across the panel using crosswalks supplied by NCES.<sup>19</sup>

Based on conversations with IPEDS-reporting institutions and the IPEDS Help Desk Staff, we identify teacher preparation programs based on the following CIP codes:<sup>20</sup>

- Education, General: 13.0100-13.0101
- Bilingual, Multilingual, and Multicultural Education 13.0201-12.0299
- Curriculum and Instruction: 13.0301
- Special Education and Teaching: 13.1000-13.1099
- Teacher Education & Professional Development, Specific Levels and Methods: 13.1200-13.1299
- Teacher Education & Professional Development, Specific Levels and Methods: 13.1300-13.1399
- Teaching English or French as a Second or Foreign Language: 13.1401-13.1499

---

<sup>19</sup> Crosswalks can be found here: <https://nces.ed.gov/ipeds/cipcode/crosswalk.aspx?y=55>

<sup>20</sup> Dan Goldhaber and Roddy Theobald provided valuable guidance here as well.

- Education, Other: 13.9999

We restrict the data to include only graduates that earned either a bachelor's or master's degree and then sum these university-level counts to the state-by-year level.

We identify teacher education program graduates in subject-specific areas using the following CIP codes:

- Elementary Education: 13.1202
- Special Education: 13.1000-13.1099
- Social Studies: 13.1317, 13.1318, 13.1332, 13.1328
- English and Language Arts: 13.1305, 13.1315
- Science Teacher Education: 13.1316, 13.1321, 13.1322, 13.1323, 13.1329, 13.1337
- Math: 13.1311

We again restrict the data to include only graduates that earned either a bachelor's or master's degree and sum the data to the state-by-year level.

## Appendix D

Our measure of union strength is a rank-based measure constructed by the Thomas B. Fordham Institute (Winkler, Scull and Zeehandelaar, 2012). Rankings are based on 21 metrics across five major areas: resources and membership, involvement in politics, scope of bargaining, state policies and perceived influence. The vast majority of data used for each metric comes from 2010 or earlier years, meaning this measure reflects the strength of unions prior to the adoption of teacher accountability reforms. We reverse the rankings such that a ranking of 1 represents the weakest union while a rank of 50 represents the strongest. We construct our measure of the obstacles to dismissing teachers in the pre-reform period using principal survey data from the 2007-2008 Schools and Staffing Survey. We estimate the proportion of principals in each state who answered yes to a set of seven binary questions asking about whether a range of factors were “barriers to the dismissal of poor-performing or incompetent teachers in their school.” We use the appropriate Schools and Staffing Survey (SASS) survey weights when constructing this measure. We then average across the seven items and rescale the resulting measure so that a one-unit change is associated with a 10 percentage point change in average perceptions of the degree to which barriers existed to dismissing poor-performing teachers. Finally, we construct two measures of the degree to which new state evaluation systems incorporate objective measures of student learning as categorized by the National Council of Teacher Quality (NCTQ, 2017). The first is a five-category ordinal measure ranging from none (1) to preponderant (5). The second is a binary indicator for objective measures of student learning being either a preponderant or significant (and explicit) percentage of teachers’ total evaluation scores.

Table D1. Testing for Differential Effects of High-Stakes Teacher Evaluation Reforms on the Number of New Teaching Licenses

|                                | (1)                 | (2)                 | (3)               | (4)                | (5)                | (6)                | (7)                | (8)                 |
|--------------------------------|---------------------|---------------------|-------------------|--------------------|--------------------|--------------------|--------------------|---------------------|
| Implement Evaluation (Eval)    | -4.193**<br>(1.463) | -4.184**<br>(1.537) | -6.520<br>(4.448) | -8.921+<br>(4.976) | -4.452+<br>(2.253) | -5.877*<br>(2.246) | -2.930*<br>(1.306) | -3.563**<br>(1.294) |
| Eval * Union Strength          | 0.063<br>(0.047)    | 0.070<br>(0.051)    |                   |                    |                    |                    |                    |                     |
| Eval * Obstacles to Dismissal  |                     |                     | 0.732<br>(0.869)  | 1.245<br>(0.970)   |                    |                    |                    |                     |
| Eval * Weight of VAM/SGP       |                     |                     |                   |                    | 0.514<br>(0.600)   | 0.982<br>(0.593)   |                    |                     |
| Eval * VAM/SGP major component |                     |                     |                   |                    |                    |                    | 0.418<br>(1.522)   | 1.875<br>(1.478)    |
| Controls                       |                     | Y                   |                   | Y                  |                    | Y                  |                    | Y                   |
| n                              | 749                 | 749                 | 749               | 749                | 749                | 749                | 749                | 749                 |

Notes: +p<.10, \*p<.05, \*\*p<.01, \*\*\*P<.001. Cells report estimates and associated standard errors clustered at the state level in parentheses. Union Strength is a continuous ranking of the strength of teacher unions across states where a higher ranking is a stronger union. Obstacles to Dismissal is a measure of the degree to which principals perceive their being obstacles to dismissing teachers. A one-unit change is equivalent to increasing the percent of principals who view there to be obstacles by ten percentage points. Weight assigned to test-based measures is an ordinal measure from 1 to 5 where higher ratings reflect state evaluation systems with larger weights assigned to measures of student learning. VAM/SGP major component is a binary indicator which takes on a value of one for states that have evaluation systems where a measures of student learning such as value-added models (VAM) or student growth percentiles (SGP) contribute the preponderate or a significant and explicit percentage of the total evaluation scores. Time-varying state-level controls are the same as those listed in Table A1. All models include state and region-by-year fixed effects.

## **Appendix E**

### **Data Sources for Concurrent Education Policy Reforms**

#### **Collective Bargaining Reforms**

We compiled data on collective bargaining reforms through a comprehensive review of state laws during our panel period. A total of six states passed legislation that eliminated mandatory collective bargaining with public teacher unions or substantially restricted the scope of which aspects of teachers' contracts were subject to collective bargaining (ID [later reversed], IN, MO, NM, TN, & WI). Ohio also made collective bargaining illegal in 2011, but the law was overturned that same year in a state-wide referendum. We control for a time-varying indicator of whether collective bargaining is not mandatory in a given state.

#### **Right-to-work laws**

We compiled data on right-to-work laws through a comprehensive review of state laws during our panel period. Four states have passed right-to-work laws that eliminated mandatory agency fees for all teachers represented by a union in collective bargaining (IN, MI, WI, & WV). West Virginia passed a right-to-work law in 2016, which was placed under a temporary injunction by the state Supreme Court. The injunction was lifted that same year after the court upheld the law. We control for a time-varying indicator of whether a state has adopted a right-to-work law.

#### **Teacher Pensions**

We measure teacher pension contributions as the required employee contribution rate, which captures the percent of total wages that teachers must contribute towards pension funds. Annual data on employee contribution rates for state and locally sponsored teacher pension funds comes from the Public Plans Data maintained by the Center for Retirement Research at Boston College. We control for a continuous measure of the percentage of teachers' salary that they contribute towards pension funds in each state and year.

#### **Teacher Licensure Test Data**

We construct indicators for different types of licensure exams states can require based on tables from the annual Digest of Education Statistics compiled by the National Center for Education Statistics (NCES). Each year the digest contains a table titled "States requiring testing for initial certification of teachers, by skills or knowledge and state." The source data for these tables come from NASDTEC (National Association of State Directors of Teacher Education and Certification) Manual on the Preparation & Certification of Educational Personnel in the United States and Canada States. We include indicators for three types of licensure exams: basic skills exams, subject exams, and professional knowledge exams. The basic skills exam is a test of fundamental reading, writing and mathematics skills. The subject exams test content knowledge in the area for which a teacher will receive licensure. The professional knowledge exam tests

knowledge of pedagogy in areas such as educational psychology, classroom management, lesson planning and evaluation and assessment.

### **Common Core State Standards**

We compiled data on the timing and duration of state's adoption of the Common Core State Standards (CCSS) from several sources including the National Council on Teacher Quality's 2015 Teacher Yearbook, the Common Core Standards Initiative Website, and an article on CCSS from Education Next. We construct a time-varying indicator for having adopted CCSS where the year refers to the fall school year.

Sources:

- [https://www.nctq.org/dmsView/Evaluation\\_Timeline\\_Brief\\_AllStates](https://www.nctq.org/dmsView/Evaluation_Timeline_Brief_AllStates)
- <http://www.corestandards.org/standards-in-your-state/>
- <http://educationnext.org/the-politics-of-common-core-assessments-parcc-smarter-balanced/>

### **Common Core Aligned States Tests**

We collected data on which states administered CCSS tests using the sources described above as well as several Education Week articles tracking which states abandoned these tests. The two test we focus on are the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balance Assessment Consortium (SBAC) tests. We control for a time-varying indicator for which states administered the PARCC or SBAC in a given year.

Additional Sources:

- [http://blogs.edweek.org/edweek/curriculum/2015/02/a\\_map\\_of\\_states\\_2015\\_testing\\_p.html](http://blogs.edweek.org/edweek/curriculum/2015/02/a_map_of_states_2015_testing_p.html)
- <https://www.edweek.org/ew/section/multimedia/state-testing-an-interactive-breakdown-of-2015-16.html>
- <https://www.edweek.org/ew/section/multimedia/states-using-parcc-or-smarter-balanced.html>