American Educational Research Journal

Math and Science Outcomes for Students of Teachers from Standard and Alternative Pathways in Texas

Michael Marder, The University of Texas at Austin

Journal:	*American Educational Research Journal*
Manuscript ID	Draft
Manuscript Type:	Original Manuscript

Abstract:	Does standard university-based preparation of STEM teachers matter? Texas provides a unique opportunity to check because it prepares more teachers through alternative certification programs than any other state. We analyze student performance on Algebra 1 and Biology exams from 2012-2016 to assess the effects of alternative and traditional teacher preparation pathways upon student learning in Texas. Students of teachers from standard programs gain around one more month of learning per year in Algebra I than students of alternatively certified teachers. Effects in Biology are weaker. Finding teacher preparation pathway differences in student learning is challenging in part because the probability a teacher is assigned more than once to a class depends on how their students performed before.

Math and Science Outcomes for Students of Teachers from Standard and Alternative Pathways in Texas

ABSTRACT

Does standard university-based preparation of STEM teachers matter? Texas provides a unique opportunity to check because it prepares more teachers through alternative certification programs than any other state. We analyze student performance on Algebra 1 and Biology exams from 2012-2016 to assess the effects of alternative and traditional teacher preparation pathways upon student learning in Texas. Students of teachers from standard programs gain around one more month of learning per year in Algebra I than students of alternatively certified teachers. Effects in Biology are weaker. Finding teacher preparation pathway differences in student learning is challenging in part because the probability a teacher is assigned more than once to a class depends on how their students performed before.

1 Introduction

TABLE 1 AROUND HERE

The United States faces persistent shortages of secondary teachers in the STEM fields. One indication of these shortages appears in Table 1. Nearly 40% of mathematics teachers either lack full teaching certification or lack a major or minor in mathematics. In the physical sciences, over 60% of teachers lack one or the other of these qualifications. Estimating from students taking Advanced Placement Computer Science (CS) exams (College Board, 2016), less than 20% of US high schools even offer computer science. The shortages may become greater because the number of teachers prepared in the highest-producing states has been falling (Figure 1).

FIGURE 1 AROUND HERE

The largest effort at the Federal level to increase the number of STEM teachers from universities is NSF's Noyce Scholarship program (National Science Foundation, 2016). In 2012, the most recent year for which data are available, around 1250 unique individuals obtained a first year of scholarship or stipend support through Noyce (Bobronnikov et al., 2014), less than 1% of the national need. Thus the institutions that have traditionally supplied the United States with teachers are not supplying enough teachers in STEM shortage areas, and programs intended to rectify these problems are doing so at a smaller scale than would be needed to solve the problem.

Teacher shortages have persisted for decades (National Commission on Excellence in Education, 1983), a fact that has led to calls to prepare teachers in new ways (Hess, 2002). The quality of teachers produced from standard pathways has also been challenged (Greenberg, McKee, & Walsh, 2013). Because of such arguments, alternative pathways to teaching now exist in all states, but there is great variation in the regulations that control what they are able or not able to do. The Every Student Succeeds Act (114th Congress, 2015) mentioned alternative certification over 30 times, and provided funding to support it, so alternative certification is likely to increase in importance.

With the increasing importance of alternative certification pathways in the United States, it is valuable to examine available evidence on the advantages and disadvantages of preparing teachers this way. In contrast to reports that paint a bleak picture of traditional teacher certification (Duncan, 2010; Greenberg et al., 2013), and in contrast to an equally negative perception that many university faculty hold of alternative certification providers (Kamnetz, 2014), the findings from the research literature are mixed. Overall, teachers prepared through alternative certification pathways are less likely to remain in the teaching profession in their early years than those coming through standard routes. Some studies find a moderate learning advantage for students whose teachers came from traditional pathways, some find a moderate advantage for students whose teachers came through alternative routes, and some are unable to discern a difference. We will review these findings in the next section.

Before stating our primary research question, it may be helpful to provide more specific information on policies with the potential to affect the supply of new STEM teachers across the nation. In the fall of 2016, the US Department of Education directed every state to develop ratings of each Teacher Preparation Program. States were required to "make meaningful differentiations in teacher preparation program performance" (US Department of Education, 2016, p. 670). As part of this, "For each year and each teacher preparation program in the State, a State must calculate the aggregate student learning outcomes of all students taught by novice teachers," where a novice teacher was "A teacher of record in the first three years of teaching" (US Department of Education, 2016, p. 656). These regulations were rescinded by Congress in the Spring of 2017, but it is plausible that the accountability metrics they required will appear again, either at the federal or at the state level.

Thus we arrive at our main research question:

> What is the effect of teachers from alternative and standard university certification pathways on high school student learning outcomes in math and science?

The setting for our study is Texas, which as shown in Figure 1 has been producing more teachers than any other state. Texas is large and varied, making it possible to access a wide range of school environments --- including urban, suburban, small town, and charter schools --- a wide degree of variation in student socio-economic status and student demographics, and a large and varied collection of teacher preparation programs. Texas presents a unique opportunity to study alternative teacher certification pathways not involving universities because as shown in Figure 1 no other state approaches Texas in the number of teachers coming from these routes. The Texas experience should also be of interest to the rest of the country because the largest Texas companies providing alternative certification are expanding to other states. As of August 2018 these companies had secured permission to operate in Florida, Nevada, Utah, Indiana, South Carolina, North Carolina, Hawaii, Arizona, Michigan, and the District of Columbia ("iTeach," 2018; "Teachers of Tomorrow," 2018).

2 Background

Value-Added Modeling

Value-added modeling arose from work of Hanushek (1971) and Sanders & Rivers (1996) and has continued to develop and improve (Koedel, Mihaly, & Rockoff, 2015; Rivkin, Hanushek, & Kain, 2005). Such models provide an important form of evidence that can be used to evaluate teacher certification pathways (Constantine, Player, Silvaa, Grider, & Deke, 2009; Guarino, Santibanez, & Daley, 2006; Wayne & Youngs, 2003).

Most studies of teacher preparation focus on a particular geographical region, either a state or a district. Well-studied regions include Florida (Harris & Sass, 2011; Sass, 2011), North Carolina (Henry, Bastian, et al., 2014; Henry, Purtell, et al., 2014), Washington State (Cowan & Goldhaber, 2016; Goldhaber, Liddle, & Theobald, 2013), Missouri (Koedel, Parsons, Podgursky, & Ehlert, 2012), New York City (Boyd et al., 2012; Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Boyd, Grossman, Lankford, Loeb, & Wykoff, 2009; Kane, Rockoff, & Staiger, 2008), California (Kane & Staiger, 2008) and Texas (Backes, Goldhaber, Cade, Sullivan, & Dodson, 2018; Mellor, Lummus-Robinson, Brinson, & Dougherty, 2008; von Hippel, Osborne, Lincove, Bellow, & Mills, 2016).

Even when student score differences between regularly and alternatively certified teachers can be discerned, they are modest compared to the scale of differences set by standard deviation on the exams. The best-studied program in the country that recruits and supports alternatively certified teachers is Teach For America (TFA) (Clark et al., 2013; Decker, Mayer, & Glazerman, 2004; Turner, Goodman, Adachi, Brite, & Decker, 2012). According to Clark et al. (2013) the difference in value-added effectiveness between TFA graduates and those of comparison programs is .06 standard deviations. These effects are measured in random controlled trials, and therefore have more internal validity than is possible from observational data. On the other hand, the schools in which the random controlled trials were conducted do not span a great range of school type so the external validity is limited.

Harris & Sass (2011) review the relationship between preservice teacher preparation and student performance. Some studies (Gordon, Kane, & Staiger, 2006; Kane et al., 2008) conclude that factors such as preparation routes and advanced degrees have almost no measurable effects on student outcomes, although there are large differences between individual teachers. Boyd et al. (2012), analyzing some of the same data from New York City as Kane et al. (2008), conclude that differences in teacher background can be detected; the difference between the studies lies in how the models were constructed. The models of Boyd et al. (2012) pay more attention to grouping together teachers with similar characteristics and from similar programs. The largest single effect in the base model of Boyd et al. (2012) is that a teacher have five years of experience, which corresponds to a value-added gain of 0.1 standard deviations in student test scores for middle school mathematics. The largest program differences, which are for Teach for America corps members, are around 0.05 standard deviations, while for College Recommended teachers the effect is around 0.02 standard deviations.

Three prior studies in Texas are particularly worth highlighting. Mellor et al. (2008) studied student learning outcomes in classrooms of novice teacher graduates from University of Texas System campuses, with data from 2003 through 2007. Their primary goal was "to determine how student achievement in the classroom might be used as an indicator of the success of teacher preparation programs" (p. 8). At the time there was no state-wide data system in place and they spent years obtaining data from over 400 districts. They carried out a variety of comparisons with multilevel models, but almost none of the effects they found was large enough to rule out having been caused by sampling uncertainty. They sum up by saying, "Our most significant finding was that limitations of most state data and assessment systems, including the one in Texas where our study was conducted, make this kind of research difficult" (p. 24). Six years later, the problem of evaluating learning gains due to Teacher Preparation Programs (TPPs) was revisited by von Hippel et al. (2016), now with the advantage of a statewide data set. They could not detect program differences and conclude, "The potential benefits of TPP accountability may be too small to balance the risk that noisy TPP estimates will encourage needless, disruptive, and ineffective policy actions" (p. 2). These conclusions are similar to the findings of Koedel et al. (2012) in Missouri, and have since been extended to several other states (von Hippel & Bellows, 2018). However Backes et al.(2018) were able to measure student learning gains for graduates of for a particular high-profile educator preparation program in Texas.

Concerning alternative certification, the National Research Council determined that "Because the information about teacher preparation and its effectiveness is so limited, high-stakes policy debates about the most effective ways to recruit, train, and retain a high-quality teacher workforce remain muddled" (NRC, 2010). Grossman & Loeb (2008, p. 185) similarly conclude that "[t]he available research does not paint a complete picture of either optimal recruiting and selection criteria nor optimal preparation opportunities." In the absence of convincing results about pathways, it is not surprising that the US Department of Education decided that "effectiveness of graduates is not associated with any particular type of preparation program, [so] the only way to determine which programs are producing more effective teachers is to link information on the performance of teachers in the classroom back to their teacher preparation programs" (US Department of Education, 2016, p. 566).

Alternative Certification in Texas

Alternative certification of teachers was first permitted in Virginia in 1982, soon followed by California, Texas, and New Jersey (Suell & Piotrowski, 2007). Alternative certification is difficult to define precisely, and can encompass a wide range of programs, but in broad terms it describes "pathways designed to attract a wider range of candidates into teaching generally by reducing or eliminating pre-service education coursework and speeding paid entry into the classroom" (Grimmett & Young, 2012).

FIGURE 2 AROUND HERE; FIGURE 3 AROUND HERE

Figure 2 shows the numbers of mathematics, science, and computer science teachers prepared in Texas since 2004 through university-based programs and non-university alternative certification. As shown in Figure 3, the total number of STEM teachers prepared in 2014-2015 was no greater than it had been in 2004-2005. The hope that alternative certification would suffice to eliminate teacher shortages has not been realized in the STEM disciplines.

TABLE 2 AROUND HERE

The National Research Council describes one particular difficulty in assessing the effectiveness of TPPs (NRC, 2010): "there is more variation within categories such as `traditional' and `alternative' –- and even within the category of master's degree programs --- than there is between the categories" (p. 2). This worry is less applicable to Texas than it may be to other jurisdictions. In practice, the regular and alternative pathways are different. Alternative certification emerged from the philosophy that barriers to teaching should be removed, and the candidates, all of whom already have finished a first Bachelor's degree, usually have a few weeks of instruction and observation after which they enter the classroom working full time, completing their pedagogical coursework during an internship year.

By contrast, standard university programs provide coursework, often but not necessarily as part of a degree, as well as fieldwork prior to a student teaching semester. As shown in Table 2, matters are not quite as simple as a binary distinction between standard and alternative programs, but cases that muddle the boundaries are not common. For example, fewer than 10% of teachers come from university post-baccalaureate programs with provisional certificates. The entities offering alternative certification programs are varied. They include school districts, state-supported education service centers and universities. However the largest providers by far are companies that advertise low cost and provide many services online ("iTeach," 2018; "Teachers of Tomorrow," 2018).

3 Sample and Methods

Overview

We present multilevel models where students are nested within classroom, classrooms are nested within teacher, teachers are nested within campus, we control for each student's pre-score, an array of demographic information about student and campus, and estimate the effect of teacher pathway on student learning. Thus, to the extent possible, the models compare teachers with other teachers teaching the same subject in the same school, and attempt to compensate for differences in school and classroom populations.

We start with data from the 2011-2012 academic year using prescores from 2010-2011 and proceed through 2015-2016. The 2011-2012 academic year was the first year that student-teacher links became available in the Texas statewide dataset, and the 2015-2016 academic year provides the most recent data available.

During our study period, Texas was transitioning between sets of high-stakes standardized exams, from TAKS to STAAR. The only high school STEM exams offered during this entire period were STAAR Algebra I and STAAR Biology. In 2011-2012, pre-scores came from TAKS 8^{th} -grade mathematics and TAKS 8^{th} –grade science, while for later years scores came from STAAR 8^{th} –grade mathematics and science.

We kept only cases where the student had a valid ninth-grade score in year *Y* and a valid eighth-grade score in the previous year. There were tens of thousands of students who took Algebra I in eighth grade, mainly high-achieving students in suburban middle schools. We decided not to include an analysis of this population in this paper. There were several accommodations available to students, including provisions for English-language learners, vision-impaired students, and a modified exam for students with learning disabilities. We could not simply group these students in with other students because most of them were taking a substantially different exam. Thus, in most of our analyses, we exclude all students who received any of these accommodations. However, we included a large subset of them in the following way. In our report on student sub-populations, we create a multilevel model for all students who took the alternate exam in year *Y-1* and also took the alternate exam in year *Y*.

FIGURE 4 AROUND HERE

Teacher years of experience is well established as an important factor in student performance (Boyd et al., 2012). It is also a confounding variable, since the probability a teacher remains in service depends upon pathway (Ramsay, 2017b). As shown in Figure 4, STEM teachers from standard programs are more likely to remain in the classroom than alternatively certified STEM teachers. Since years of service is a consequence of preparation pathway, one should not control for it (Morgan & Winship, 2015, pp. 101–109). That does not mean years of experience can completely be ignored. For example, the 2016 teacher preparation regulations mentioned earlier specified that teacher preparation programs were to be assessed on the performance of their novice teachers, those with less than 4 years of experience (US Department of Education, 2016, p. 68), and this choice is under consideration by states as well. Thus, we decided to use two teacher experience groups: novice teachers, defined as those with less than four years of experience, and all teachers with up to twenty years of experience. It is only during the last two decades that the percentage of teachers alternatively certified in a year moved above 10%, and online and for-profit certification routes came into existence. Thus this is the largest population of teachers containing information about the pathways we are investigating. Table 3 indicates the numbers of teachers in our sample for each academic year and level of experience.

TABLE 3 AROUND HERE

To deal with the pathway complexity portrayed in Table 2, in forming comparison groups we decided to say that teachers prepared by Institutions of Higher Education (IHE) enrolled in standard or post-baccalaureate programs and obtaining standard first certificates came from a Standard Program. Everyone else, including some of the students from universities, we attributed to an Alternative Program. We ran variants of the analysis, for example including graduates of university-based Alternative programs who began teaching with a standard certificate in the Standard group. However as the numbers of teachers of uncertain classifications were small, and none of the conclusions in our analysis were affected, we report only results from the comparison groups we have just described. Note that the great majority of teachers come either from standard university programs with student teaching or from alternative programs without it.

TABLE 4 AROUND HERE

Sample Characteristics

Table 4 provides descriptions of students appearing in our sample, comparing the classrooms of teachers from standard and alternative pathways. We aggregate all years together, as changes over time were not worth remarking. Standard and alternatively certified teachers have significantly different student populations. The alternatively certified teachers have a higher fraction of their students who are eligible for free and reduced lunch and who are Black and Hispanic. In general, for any factor that tends to lead to lower student outcomes, alternatively certified teachers have more of these students. This makes it important to control for these student characteristics in the analysis.

We investigated whether the difference in student populations of the teachers from the two pathways is mainly within schools or between schools. To do this we constructed hierarchical linear models (Bolker et al., 2016) of the form

$$X_i \sim N(C_i + Cert_i; \sigma_S^2);\ C_i \sim N(\mu_C; \sigma_C^2)$$

where X_i is a demographic characteristic of student i, C_i is their campus, and $Cert_i$ is the certification status of their teacher. The results appear in the final column of Table 4. They show that after one controls for the campus, the difference between student populations of teachers from standard and alternative pathways becomes much smaller, often insignificant, and may reverse sign. Therefore the difference in student populations is almost completely due to the schools in which teachers from standard and alternative pathways are likely to work. Within a given school the populations to which they are assigned are much less easily distinguished.

Months of Schooling Units (MOS)

Test score effect sizes are customarily obtained by dividing exam scores by the standard deviation. For ninth graders who took Algebra I in 2011-2012, the standard deviation was 0.17. For these same students the standard deviation of their mathematics scores the year before in 8th grade was 0.15, and other years are similar. To set the units, we employ 0.16 as the standard deviation. We define Months of Schooling units where 9 Months of Schooling (MOS) corresponds to one quarter of a standard deviation (Gates Foundation, 2012). In these units, gaining 0.04 in raw score on an exam, is reported as 9 Months of Schooling.

Weights

Any given student test score result could end up in our data set from one to six times. The test scores appeared multiple times when the student took classes with separate identification numbers in separate semesters, when more than one teacher was associated with the class section, or if the student changed schools in Texas. We weighted every student record inversely with the number of times they appeared, so if a student was taught by several teachers during the year, each of them shared equally, and that student did not contribute more to the final results than a student who appeared only once.

Multilevel Models

We explored many different multilevel models, using lmer in R (Bolker et al., 2016). In our first collection of models, Eq. (1), all data from 2011-12 through 2015-16 were included at once. At the top level, $\mathrm{S_i}$ is the score of student i in some year and $\mathrm{S_{i,\nu-1}}$ is the same student's score on the exam in the same subject the previous year in a cubic polynomial. Teacher j of student i contributes through the random intercept $T_{j[i]}$. By modeling the teacher in this way, each teacher should contribute equally to the estimate of the effect of their pathway to teaching (Koedel, Parsons, Podgursky, & Ehlert, 2015). The campus k contributes random intercept $C_{k[i]}$ as does the class n through $\mathrm{Class}_{n[i]}$. Coefficients for student-level demographic factors X range over Gifted, racial and ethnic groups, Limited English Proficiency (LEP), Free/Reduced Lunch Eligibility (EcoDis), and Special Education. Here g[i] is the value of group membership for student i. We modeled the influence of tracking (Jackson, 2014). The most important form of student tracking in Texas is placing students in Algebra I in 8th grade. We removed the 8th grade population from our study of mathematics and created a flag for those who had Algebra I in 8th grade when modeling biology. We control for classroom-level averages of the demographic variables through $\gamma_X\overline{X}_{n[i]}$ (Chetty, Friedman, & Rockoff, 2014; Friedman, Rockoff, & Chetty, 2014). We included a variable E to control for teacher years of experience, assuming binned values of 0-4, 4-10, and 10-20 years of experience

The main item of interest, certification pathway Cert_m of teacher j of student i enters as a fixed effect. Finally, the second level of the model has random intercepts for teacher T, campus C, and class section Class,

$$S_{i,t} \sim N\left(\sum_{\nu=2012}^{2016} \sum_{\beta=1}^{3} \lambda_{\beta\nu} S_{i,\nu-1}^{\beta} + T_{j[i]} + E_{j[i]} + C_{k[i]} + \mathrm{Class}_{n[[i]]} + \mathrm{Cert}_{m[j[i]]} + \sum_X X_{g[i]} + \sum_X \gamma_X \overline{X}_{n[i]}; \sigma_S^2\right). \quad (1)$$

$$T_j \sim N(\mu_T; \sigma_T^2); \quad C_k \sim N(\mu_C; \sigma_C^2); \quad \mathrm{Class}_n \sim N(\mu_L; \sigma_L^2) \quad .$$

Tables 5 and 6 provide results from these models. We present results from four different variants of this general form so as to display the effect upon the variable of interest, Cert, of progressively adding terms. Model (1a) lacks a campus intercept and also lacks averages of classroom demographics. The argument for this model would be that strong campuses with privileged students are strong mainly because their teachers are strong. Model (1b) adds the campus intercept as a random effect, (1c) adds classroom averages of demographic variables, and (1d) adds a control for years of service. Model (1c) is probably the most persuasive of the model specifications, since classroom averages of demographic variables do turn out to affect the results, while controlling for years of service in (1d) is debatable on causal grounds, but adding it or not does not make much of a difference.

We also ran models in which each year was treated separately. Model 2 is similar to model (1c) with random intercepts for campus, class, and teacher, and controls both for student demographics, classroom averages of student demographics, and tracking.

$$S_{i,t} \sim N(\textstyle\sum_{\beta=1}^{3} \lambda_\beta\, S_{i,t-1}^{\beta} + T_{j[i]} + C_{k[i]} + \mathrm{Class}_{n[[i]]} + \mathrm{Cert}_{m[j[i]]} + \sum_X X_{g[i]} + \sum_X \gamma_X \overline{X}_{n[i]}\,;\, \sigma_S^2). \quad (2, \text{Random})$$

$$T_j \sim N(\mu_T; \sigma_T^2); \quad C_k \sim N(\mu_C; \sigma_C^2); \quad \mathrm{Class}_n \sim N(\mu_L; \sigma_L^2) \quad .$$

Model 3 has the form

$$S_{i,t} \sim N(\textstyle\sum_{\beta=1}^{3} \lambda_\beta\, S_{i,t-1}^{\beta} + T_{J[I]} + C_{k[i]} + \mathrm{Class}_{n[j[i]]} + \mathrm{Cert}_{m[j[i]]} + \sum_X X_{g[i]} + \sum_X \gamma_X \overline{X}_{n[i]}; \sigma_S^2) \quad (3, \text{Fixed})$$

$$T_j \sim N(\mu_T; \sigma_T^2);\ \mathrm{Class}_n \sim N(\mu_L; \sigma_L^2)$$

This is the same as the previous, except that campus is treated as a fixed effect at the top level, rather than being modeled as a random effect at the second level. This model is less appropriate for finding the contribution of teacher pathway because in cases where a campus has teachers from only a single pathway, the campus fixed effect subtracts them off rather than comparing them with teachers in similar campuses as the campus random effect model does.

Model 4 is

$$S_{i,t} \sim N(\textstyle\sum_{\beta=1}^{3} \lambda_\beta\, S_{i,t-1}^{\beta} + T_{J[I]} + \mathrm{Class}_{n[j[i]]} + \mathrm{Cert}_{m[j[i]]} + \sum_X X_{g[i]} + \sum_X \gamma_X \overline{X}_{n[i]}\,;\, \sigma_S^2) \quad (4, \text{None})$$

$$T_j \sim N(\mu_T; \sigma_T^2);\ \mathrm{Class}_n \sim N(\mu_L; \sigma_L^2)$$

In this case there is no campus intercept. One would use this model if one adopts the view that the difference between school performance is mainly due to the teachers they get and not to other non-student factors.

We had a fifth model which we applied to subgroups of students.

$$S_{i,t} \sim N(\textstyle\sum_{\beta=1}^{3} \lambda_\beta S_{i,t-1}^{\beta} + T_{J[I]} + \text{Class}_{n[j[i]]} + \text{Cert}_{m[j[i]]} + \sum_X \gamma_X \overline{X}_{n[i]}\ ;\ \sigma_S^2) \qquad \text{(5, Subgroup)}$$

$$T_j \sim N(\mu_T; \sigma_T^2); \quad C_k \sim N(\mu_C; \sigma_C^2); \quad \text{Class}_n \sim N(\mu_L; \sigma_L^2)$$

This model was applied after being restricted to a demographic subset of our sample, for example to the subgroup of economically disadvantaged students, gifted students, or students taking an alternate test. This model allowed us to focus on the effect of teacher pathway on a specific demographic of students, while also controlling for the broader demographic compositions of their classes.

5 Student Learning

Overall Results

TABLE 5 AROUND HERE

TABLE 6 AROUND HERE

We begin our discussion of results with Model (1). The random and fixed effects are given in Table 5 for Algebra I and Table 6 for Biology. The coefficients related to student subgroups for the models of Algebra I and Biology are quite similar to each other.

Among the random effects, the largest is the difference between campuses, with a standard deviation of 9 Months of Schooling for Algebra I and 7 in Biology. The standard deviation of the difference between teachers is around 8 Months of Schooling in Algebra I and 6 Months in Biology. That is, we find slightly larger differences between campuses than within them. The standard deviation of classes taught by the same teacher is between 5 and 6 Months of Schooling in both subjects.

We report the teacher pathway effect (Cert) as positive when students of teachers with standard certification get higher scores than those of alternatively certified teachers. Overall results appear in Tables 5 and 6. The students of standard certified teachers gain around 1 more month of schooling per year in Algebra I. This result is significant in all the model specifications chosen, although the effect is larger in the models with fewer controls. That is not surprising, since we know from Table 4 that teachers from standard programs find themselves overall in classrooms with fewer economically disadvantaged students, and more white students. The question we cannot fully settle here is the degree to which the performance of economically disadvantaged students is due to their teachers. In Biology, results are significant in models 1a and 1b , but not in models 1c and 1d that include simultaneously averages of classroom demographics and an intercept for each campus. Thus the case for improved student learning in classrooms with teachers from standard programs is weaker in Biology than in Algebra I, as it only survives in models that attribute most of the student learning gains in classrooms of privileged students to the teachers rather than the concentration of privilege.

The difference between models 1c and 1d is that 1d includes a control for years of experience, and 1c does not. Controlling for years of experience is not clearly a bad choice since the population of teachers with more than 10 years of experience has not yet reached steady state in the contribution of teachers from alternative pathways. Alternative certification pathways were still growing rapidly in size when teachers with 20 years of experience entered the profession. This makes it unclear how years of experience should best be treated. Fortunately, one sees that whether it is included or not turns out not to make much of a difference.

We now turn to a more detailed analysis year by year and for various subgroups. Table 7 presents results for the three different models that vary the way campus effects are treated. The fixed effect model (Eq. 2, Fixed) which maximizes how much of an effect is attributed to the campus tends to give the smallest results, and the model without campus effects (Eq. 3, None) tends to give the largest results, showing that strong Algebra I teachers are associated with strong campuses. In every year and for each of the models, students gain about one month more learning per year in classes of Algebra I teachers from standard programs with up to 20 years of experience. The point estimates in classrooms of novice Algebra I teachers with up to 4 years of experience are similar, but the sample sizes are much smaller, and few of the differences are statistically significant. For Biology classrooms, there are almost no statistically significant results, and the point estimates are scattered between positive and negative. If one removes the classroom averages of demographic variables from models (2)-(5), there are many cases where students Biology teachers from standard programs have significantly higher learning gains, but these models are harder to defend than the ones we have used, and we do not report their results.

TABLE 7 AROUND HERE

Tables 8 and 9 provide results for each combination of teacher pathway, years of experience, subject, and a variety of student subgroups, using the model in Eq. 5. Almost all of the estimates in Algebra I by subgroup indicate that students of teachers with standard certification gain between one and four more Months of Schooling per year more than their counterparts whose teachers were alternatively certified. The largest differences are for students flagged as gifted, but the results from students eligible for free and reduced lunch (FRL) and those of limited English proficiency (LEP) are also noteworthy. The effects are stronger in Algebra I for teachers with up to 20 years of experience than they are for novice teachers. In Biology there are almost no statistically significant results, except for novice teachers in 2011-2012. For both Algebra I and Biology the majority of the point estimates favor teachers from standard programs, and there is only one statistically significant result favoring teachers from alternative pathways, which is for LEP students of novice Biology teachers in 2015-2016.

The Algebra I results from this section are summarized in graphical form in Figure 5. The results for all students come from the model in Eq. (2), and the rest from Eq. (5). Only results for teachers with up to 20 years of experience are shown in the graph.

TABLE 8 AROUND HERE

TABLE 9 AROUND HERE

FIGURE 5 AROUND HERE

6 Teacher Assignment

We wondered whether our results might be affected by the way teachers were selected to teach courses with high-stakes exams. Because of the very high stakes for schools and their personnel associated with these exams, one could expect principals to monitor past results carefully, and assign teachers with a good track record for raising student test scores to Algebra I and Biology (Dieterle, Guarino, Reckase, & Wooldridge, 2015).

We find evidence for such assignment bias, and it shows up in several ways. We constructed a two-stage model for the probability of being assigned to teach. The first stage of the model is Equation 1, which computes a value-added coefficient T for every teacher. The second stage is a binomial logistic regression model that computes the probability a teacher was assigned to teach Algebra I or Biology as a function of the value-added score in the course in the same school the year before. Thus, the probability of being assigned ($a = 1$) to a course given value-added score T in the previous year and certification pathway Cert is

$$P(a|T,\text{Cert}) = 1/(1 + \exp[-(a_0 + a_T T + a_c \text{Cert})]) \ . \quad (5)$$

Here T is the value-added score we compute for each teacher normalized by the standard deviation of value-added scores and Cert=1 corresponds to a teacher who came from a standard program. The coefficients of this model for Algebra I and Biology appear in Table 10.

TABLE 10 AROUND HERE

The results are significant every year. For example, if an Algebra I teacher from a standard program in 2011-2012 had a value-added score 1.5 standard deviations above the mean they had a 60% chance of returning to teach the course the next year as opposed to a 41% chance if their value-added score was 1.5 standard deviations below the mean.

School principals and department heads did not of course have access to the specific value-added scores we have computed, but they had in their possession all the raw data about student achievement that go into making them up and appear to have acted accordingly (Dieterle et al., 2015; Grissom, Loeb, & Nakashima, 2014).

FIGURE 6 AROUND HERE

While value-added score was the strongest single predictor we found of whether a teacher was assigned twice in a row to teach Algebra I or Biology, many characteristics of the teacher population changed. It is worth nothing that the STAAR exams we use as a post-test were employed for the first time in the spring of 2012; at that time high school students were expected to take 15 exams in order to graduate. In the spring of 2013, the Texas legislature reduced the required exams from 15 to 5 and abolished all the STEM exams but Algebra I and Biology. One result was a dramatic shift in the age distribution of teachers assigned to Algebra I and Biology. In 2011-2012, 35% of all Texas mathematics teachers had 0-5 years of experience (Ramsay, 2017a), and 39% of the Algebra I teachers had 0-5 years of experience. But by 2014-2015, when the percentage of Texas mathematics teachers with 0-5 years of experience was essentially unchanged at 37%, the percentage of Algebra I teachers with 0-5 years of experience dropped to 17%. As shown in Figure 8, the drop in novice Algebra I teachers was accompanied by a rise in teachers with 10-20 years of experience. Placement of biology teachers is equally well predicted by value-added scores, and the distribution of Biology teachers changed in a very similar fashion, from a distribution characteristic of science teachers overall, to a distribution greatly weighted towards teachers with 10 to 20 years of experience.

7 Conclusions

Prior studies have concluded that characteristics of teacher education are too small to detect or too small to matter in student achievement (Aaronson, Barrow, & Sander, 2007; Gordon et al., 2006; Harris & Sass, 2011; Rivkin et al., 2005; Staiger & Rockoff, 2010; von Hippel & Bellows, 2018). We found significant effects on the order of one Month of Schooling for ninth-grade students of Algebra I in favor of teachers with standard certification. However effects in Biology are weaker and do not clearly favor one pathway or another in the models with the strongest controls. The pathway effects we find in Tables 5 through 9 are small compared to the standard deviation of teachers or schools, but they are consistent with teacher pathway effects found in other studies (Boyd et al., 2009).

FIGURE 7 AROUND HERE

Whether gaining one Month of Schooling is an important or unimportant educational difference merits additional discussion. It corresponds to 2.7% in standard deviation units, or a 20% greater chance of getting one more problem right on a 50-question exam. This may seem too small to matter. However, if sustained over time, it is an effect comparable to that of living in poverty. Figure 7, building on techniques of Author (2012) shows that if one groups students according to their mathematics scores and free/reduced lunch status in 4th grade, and then follows the students through 11th grade, the difference between the well-off and low-income students develops to around 3 Months of Schooling, and it takes around 3 years to develop. That is, the difference in test score results due to having a math teacher from a standard program for a year is of the same order as the effect on test results over a year associated with living in poverty. One could conclude that the standardized tests are not very sensitive either to instruction (Popham, 2007; Stroup, 2009) or to poverty, but this does not mean the tests are completely incapable of detecting them.

Because the tests are not very sensitive to what we wish to measure, it takes a large number of teachers and students to arrive at reliable values. As seen in Tables 5 and 6, for a single teacher teaching multiple sections of the same class at the same time, the typical variation from one section of the class to another is around 5 Months of Schooling. We estimate that to find the effect of any particular type of teacher preparation pathway with uncertainty less than 1 Month of Schooling, one must average results from around 1000 teachers. We accomplished this by aggregating together preparation programs in groups with similar practices rather than focusing on effects at the level of a single program. Such grouping also is a feature of the finding of positive associations between National Board certification and student achievement in Cowan and Goldhaber (2016), and of pathways in New York City by Boyd et al. (2009).

We found overall that Algebra I teachers from standard certification pathways improved student test scores by around 1 Month of Schooling. For subgroups including gifted students, students eligible for free and reduced lunch, and Black and Hispanic students, students of standard teachers gained 1.5 to 3 months more schooling. In Biology the evidence for positive effects for teachers from standard programs overall is not robust, except that in 2011-2012, for students eligible taking the alternative exam, students with limited English proficiency, and Black and Hispanic students, novice Biology teachers from standard programs added 1.5-4 more Months of Schooling than alternatively certified novice teachers. Tables 8 and 9 provide many cases where in a particular year and with a particular group, teachers from standard programs obtained significantly higher student gains than teachers from alternative programs, while there is only one statistically significant instance of the reverse effect.

It has frequently been stated that that variance between classrooms within schools is larger than variance between schools (Nye, Konstantopoulos, & Hedges, 2004). Staiger & Rockoff (2010) conclude that "School leaders have very little ability to select effective teachers during the initial hiring process" and present as evidence "the fact that most of the variation in teacher effects occurs among teachers hired into the same school." Our results are different. In Tables 5 and 6, variation between schools is the largest random effect, followed by variation between teachers in schools, then followed by different classrooms of the same teacher. Thus if we take into account both the variance between schools and the way we found teachers to be assigned, evidence indicates that principals hire teachers and assign them to classes based on information about their effectiveness.

We obtain a result, often found before, that there is more variation of student outcomes within teacher preparation pathways than between preparation pathways. This finding has been used in support of policies that reduce barriers for new people to enter teaching, but make it difficult for them to continue unless they can demonstrate favorable student outcomes (Gordon et al., 2006). While such policies might make sense in cases where there are more people wishing to become teachers than there are positions available, they are less justifiable for shortage areas such as secondary STEM. It is hard to imagine that young people or career changers will be attracted to secondary teaching by the prospect of high-stakes evaluations from multilevel models operating on their students' test scores. Newspaper accounts such as those of Bonner (2016) conclude that these evaluations have been exacerbating teacher shortages. Teacher shortages may not directly impact high-stakes subjects such as Algebra I and Biology; schools have to staff them or face severe penalties. Shortages show up for subjects such as computer science where there are no high-stakes assessments, and where only a small fraction of high schools even offer a course (Guzdial, 2012). It is tempting to consider policies that make it difficult for teachers with low value-added scores to continue teaching altogether, in hopes of capturing some of the 7 Months of Schooling advantage for the best teachers in Tables 5 and 6. However, reducing the stability of teaching careers will impact which individuals decide to enter teaching or settle instead on other careers. There is no assurance that secondary students will benefit in the end.

Our results do not justify the disruptions that would come from an abrupt policy change impacting alternative certification programs in Texas. In Algebra I there are many subgroups of students for whom the advantages of having a teacher from a standard university program are significant, while in Biology the effects are weaker. In view of the substantial extra time students spend preparing in standard pathways before full-time teaching, universities should consider whether there are any lessons to learn from the alternative programs. One must keep in mind because of the shortage of STEM teachers that it is difficult to justify reducing teachers from any pathway. Slightly increasing the scores of low-income students on Algebra exams but reducing the number able to take Physics or Chemistry at all would almost certainly be a very poor trade. On the other hand, our results do not provide strong incentive for other states to follow Texas's lead in establishing a large for-profit alternative certification sector. As shown in Figures 2 and 3, the growth of alternative certification in Texas since the mid 2000s has not led in the end to an increase in the production of STEM teachers.

The change we found in teacher assignment over time provides additional reason to worry about how high-stakes tests are being used. The tests were designed to measure student mastery of academic material. They are now being used to allow students to advance academically, to judge the performance of individual teachers, to judge the performance of schools and school administrators, and finally to judge the programs that prepare the teachers. For test results to provide an unbiased estimate of preparation programs, principals would have to ignore teachers' track record when assigning them to classes with high stakes assessments, even when the future of both students and the administrators are at risk, and when administrators are constantly impressed with the importance of data-driven decision making (Houston, 2013). This is not realistic.

The literature on teacher preparation pathways has paid little attention to how teachers were assigned. In experimental studies, students were randomized between classes, but the teachers were not. As we have shown here, teacher assignment can have a large effect on measurement of preparation program effectiveness.

We have obtained robust evidence that Algebra I students learn more when their teachers come from standard programs. The learning gains are most pronounced for groups such as gifted, Hispanic, and those eligible for free and reduced lunch. For Biology the situation is less certain. Students of teachers from standard programs gain around one more month of learning in models that emphasize the importance of teachers in the accomplishments of privileged students, but in models constructed to attribute the accomplishments of privileged students to factors other than the teacher, differences between standard and alternative certification largely disappear.

Around 700,000 undergraduates obtain STEM degrees from US universities each year. This is an enormous pool; persuading just 1% more to obtain a teaching certificate along with their degree each year would add 7000 new STEM teachers. Thus, we encourage support for the preparation of STEM teachers through standard university pathways as the most efficient, scalable, and high-quality way to address the critical need for improved STEM education and to address the shortage of STEM teachers.

Bibliography

114th Congress. (2015, December). Every Student Succeeds Act. Retrieved from https://www.congress.gov/bill/114th-congress/senate-bill/1177/text

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135.

Backes, B., Goldhaber, D., Cade, W., Sullivan, K., & Dodson, M. (2018). Can UTeach? Assessing the relative effectiveness of STEM teachers. *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2018.05.002

Bobronnikov, E., Price, C., Gamse, B., Parsad, A., Roy, R., & Velez, M. (2014). Preliminary Findings from the Noyce Program Evaluation and Possible Future Directions, Unpublished presentation at 2014 Noyce Conference.

Bolker, B., Bates, D., Maechler, M., Walker, S., Christensen, R. H. B., Singmann, H., … Green, P. (2016). Package lme4. Retrieved from https://cran.r-project.org/web/packages/lme4/lme4.pdf

Bonner, L. (2016). Enrollment plunges at UNC teacher prep programs. Retrieved from http://www.charlotteobserver.com/news/local/education/article58311793.html

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The Effect of Certification and Preparation on Teacher Quality. *Future of Children*, *17*(1), 45–68.

Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2012). Recruiting Effective Math Teachers: Evidence from New York City. *American*

Educational Research Journal, *49*(6), 1008–1047.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wykoff, J. (2009). Teacher Preparation and Student Achievement. *Educational Evaluation and Policy Analysis*, *31*, 416–440.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Clark, M. A., Chiang, H. S., Silva, T., McConnell, S., Puma, K. S. A. E. M., Warner, E., & Schmidt, S. (2013). The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs. Retrieved from http://ies.ed.gov/ncee/pubs/20134015/

College Board. (2016). AP Program Participation and Performance Data 2016. Retrieved from https://research.collegeboard.org/programs/ap/data/participation/ap-2016

Constantine, J., Player, D., Silvaa, T., Grider, K. H. M., & Deke, J. (2009). An evaluation of teachers trained through different routes to certification, final report. Retrieved from ies.ed.gov/ncee/pubs/20094043/pdf/20094043.pdf

Cowan, J., & Goldhaber, D. (2016). National Board Certification and Teacher Effectiveness: Evidence From Washington State. *Journal of Research on Educational Effectiveness*, *9*(3), 233–258. https://doi.org/10.1080/19345747.2015.1099768

Decker, P. T., Mayer, D. P., & Glazerman, S. (2004). The effects of Teach For America on students: Findings from a national evaluation. Retrieved from https://www.mathematica-mpr.com/-/media/publications/pdfs/teach.pdf

Dieterle, S., Guarino, C. M., Reckase, M. D., & Wooldridge, J. M. (2015). How do Principals Assign Students to Teachers? Finding Evidence in Administrative Data and the Implications for Value Added. *Journal of Policy Analysis and Management*, *34*, 32–58. Retrieved from http://www.scopus.com/inward/record.url?scp=84916210067&partnerID=8YFLogxK

Duncan, A. (2010). Teacher Preparation: Reforming the Uncertain Profession. Retrieved from http://www.ed.gov/news/speeches/teacher-preparation-reforming-uncertain-profession

Friedman, J. N., Rockoff, J. E., & Chetty, R. (2014). Measuring the Impacts of Teachers I : *American Economic Review*, *104*(9), 2593–2632. https://doi.org/10.3386/w19423

Gates Foundation. (2012). Gathering Feedback for Teaching: Combining high quality observations with student surveys and achievement gains. Research Paper. Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review*, *34*, 29–44. https://doi.org/10.1016/j.econedurev.2013.01.011

Gordon, R., Kane, T. J., & Staiger, D. O. (2006). Identifying Effective Teachers Using Performance on the Job. Retrieved from www.brookings.edu/views/papers/200604hamilton_1.pdf

Greenberg, J., McKee, A., & Walsh, K. (2013). Teacher Prep Review: A review of the nation's teacher preparation programs. Retrieved from http://www.nctq.org/dmsView/Teacher_Prep_Review_2013_Report

Grimmett, P. P., & Young, J. C. (2012). *Teacher Certification and the Professional Status of Teaching in North America: The new battleground for public education*. Charlotte, NC: Information Age Publishing.

Grissom, J. A., Loeb, S., & Nakashima, N. A. (2014). Strategic Involuntary Teacher Transfers and Teacher Performance: Examining Equity and Efficiency. *Journal of Policy Analysis and Management*, *33*(1), 112–140. https://doi.org/10.1002/pam.21732

Grossman, P., & Loeb, S. (2008). *Alternative Routes to Teaching: Mapping the New Landscape of Teacher Education.* Harvard Education Press.

Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher recruitment and retention: a review of the recent empirical literature. *Review of Educational Research*, *76*(2), 173–208. https://doi.org/10.3102/00346543076002173

Guzdial, M. (2012). Why isn't there more computer science in US high schools? Retrieved from http://cacm.acm.org/blogs/blog-cacm/156531-why-isnt-there-more-computer-science-in-u-s-high-schools/fulltext

Hanushek, E. (1971). Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data. *American Economic Review*, *61*, 280–288.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, *95*(7–8), 798–812. https://doi.org/10.1016/j.jpubeco.2010.11.009

Henry, G. T., Bastian, K. C., Fortner, C. K., Kershaw, D. C., Purtell, K. M., Thompson, C. L., & Zulli, R. A. (2014). Teacher Preparation Policies and Their Effects On Student Achievement. *Education Finance and Policy*, *9*(3), 264–303. https://doi.org/10.1162/EDFP_a_00134

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The Effects of Teacher Entry Portals on Student Achievement. *Journal of Teacher Education*, *65*(1), 7–23. https://doi.org/10.1177/0022487113503871

Hess, F. M. (2002). Tear Down This Wall: The Case for a radical overhaul of teacher certification. *Educational Horizons, 80*, 169–183. Retrieved from http://www.jstor.org/stable/42927125

Houston, P. (2013). *Using Data to Improve Schools: What's Working*. Retrieved from http://www.aasa.org/uploadedFiles/Policy_and_Advocacy/files/UsingDataToImproveSchools.pdf

iTeach. (2018). Retrieved July 1, 2018, from http://iteach.net

Jackson, C. K. (2014). Teacher Quality at the High School Level: The Importance of Accounting for Tracks. *Journal of Labor Economics*, *32*, 645–684.

Kamnetz, A. (2014). For teachers many paths into the classroom...some say too many. Retrieved from http://www.npr.org/sections/ed/2014/09/12/347375798/for-teachers-many-paths-into-the-classroom-some-say-too-many

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615–631.

Kane, T. J., & Staiger, D. O. (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. Retrieved from http://www.nber.org/papers/w14607

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-Added Modeling: A Review. *Economics of Education Review*, *47*, 180–195.

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2012). Teacher preparation programs and teacher quality: Are there real differences across programs? Retrieved from http://www.caldercenter.org/sites/default/files/WP-79.pdf

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2015). Teacher Preparation Programs and Teacher Quality: Are There Real Differences Across Programs? *Education Finance and Policy*, *10*, 508–534.

Mellor, L., Lummus-Robinson, M., Brinson, V., & Dougherty, C. (2008). Linking Teacher Preparation Programs to Student Achievement, unpublished report, available on request from corresponding author of this article.

Morgan, S., & Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (Second Edi). Cambridge University Press.

National Commission on Excellence in Education. (1983). A Nation at Risk. Retrieved from http://www2.ed.gov/pubs/NatAtRisk/index.html

National Science Foundation. (2016). Robert Noyce Teacher Scholarship Program. Retrieved from http://www.nsf.gov/pubs/2016/nsf16559/nsf16559.pdf

NRC. (2010). *Preparing Teachers: Building Evidence for Sound Policy*. National Research Council.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*, 237–257.

Popham, W. J. (2007). Instructional insensitivity of tests: Accountability's dire drawback. *Phi Delta Kappan*. https://doi.org/10.1177/003172170708900211

Ramsay, M. (2017a). *Experience of Math and Science Teachers 2012-2016*. Retrieved from http://tea.texas.gov/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=51539614334&libID=51539614334

Ramsay, M. (2017b). *Teacher Retention by Route 2012-2016*. Retrieved from http://tea.texas.gov/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=51539614330&libID=51539614330

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*, 417–458.

Sanders W.L., & R. J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement: Research progress report.* Knoxville, TN: University of Tennessee.

Sass, T. R. (2011). Certification Requirements and Teacher Quality: A Comparison of Alternative Routes to Teaching. Working Paper 64. *National Center for Analysis of Longitudinal Data in Education Research*.

Staiger, D. O., & Rockoff, J. E. (2010). Searching for Effective Teachers with Imperfect Information. *Journal of Economic Perspectives*, *24*(3), 97–117.

https://doi.org/10.1257/jep.24.3.97

Stroup, W. M. (2009). What it means for mathematics tests to be insensitive to instruction. In *Plenary Address at the Thirty-First Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*.

Suell, J. L., & Piotrowski, C. (2007). Alternative teacher education programs: a review of the literature and outcome studies. *Journal of Instructional Psychology*, *34*(1), 54–58.

Teachers of Tomorrow. (2018). Retrieved July 19, 2018, from http://teachersoftomorrow.org

Turner, H. M., Goodman, D., Adachi, E., Brite, J., & Decker, L. E. (2012). *Evaluation of Teach For America in Texas Schools*. Retrieved from http://scee.groupsite.com/uploads/files/x/000/093/27d/Evaluation of Teach for America in Texas Schools.pdf?1366060957

US Department of Education. (2016). Teacher Preparation Issues: Final regulations. Retrieved from http://www2.ed.gov/documents/teaching/teacher-prep-final-regs.pdf

von Hippel, P. T., & Bellows, L. (2018). How much does teacher quality vary across teacher preparation programs? Reanalyses from six states. *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2018.01.005

von Hippel, P. T., Osborne, C., Lincove, J., Bellow, L., & Mills, N. (2016). Teacher Quality Differences Between Teacher Preparation Programs: How Big? How Reliable? Which Programs Are Different? *Economics of Education Review*, *53*, 31–45. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2506935

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122. https://doi.org/10.3102/00346543073001089

FIGURE CAPTIONS

Figure 1: Teachers prepared by state over 6-year period (US Department of Education, 2018, Completers 2008-2016). IHE refers to Institution of Higher Education.

Figure 2: STEM teacher production in Texas from 2004 until 2017, comparing production from regular and alternative certification pathways. Data from Texas Education Agency.

Figure 3: STEM teacher production in Texas from 2004 until 2017, summing all pathways. Data from Texas Education Agency.

Figure 4: Retention of Texas STEM teachers in teaching by preparation pathway, averaged over cohorts entering from 2004 until 2013. Data from Texas Education Agency through Education Research Center.

Figure 5: Added months of schooling for all Algebra I teachers from standard programs as compared with teachers from alternative programs, overall and for student subgroups, showing change over time. Data from Texas Education Agency through Education Research Center.

Figure 6: Distribution of teacher years of experience in Algebra I and Biology. Data from Texas Education Agency through Education Research Center.

Figure 7: Mathematics exam averages in Months of Schooling units over time for cohorts of students from the Class of 2011 grouped both by their scores in 4^{th} grade, and by their free/reduced lunch status. Data from Texas Education Agency through Education Research Center.

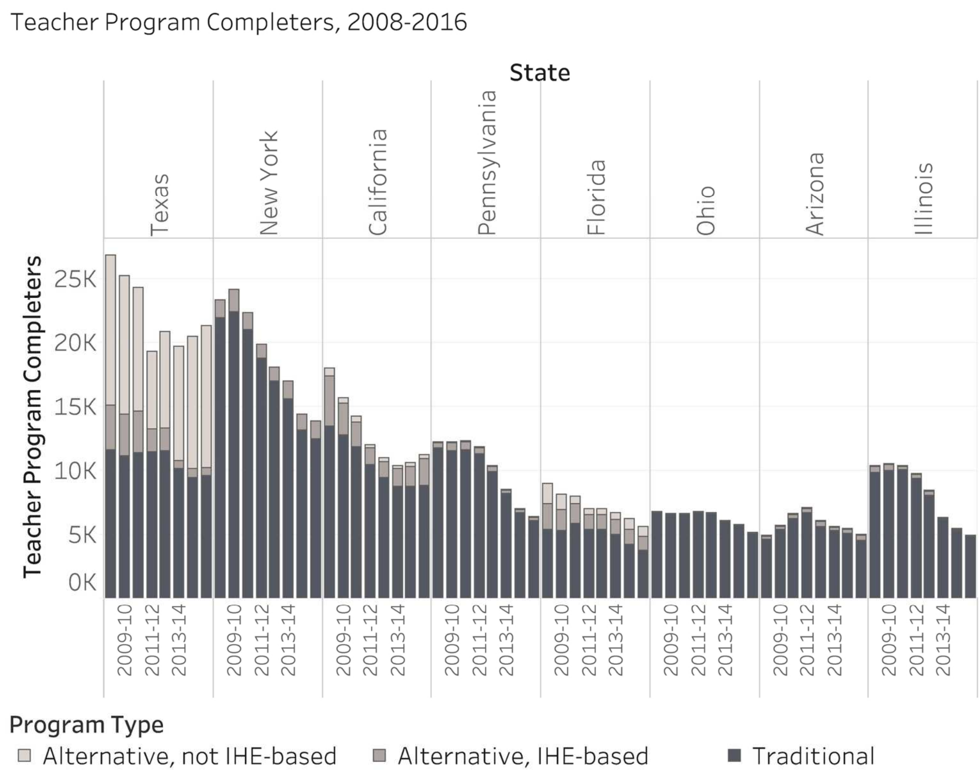

Figure 1: Teachers prepared by state over 6-year period (US Department of Education, 2018, Completers 2008-2016). IHE refers to Institution of Higher Education.

158x129mm (300 x 300 DPI)

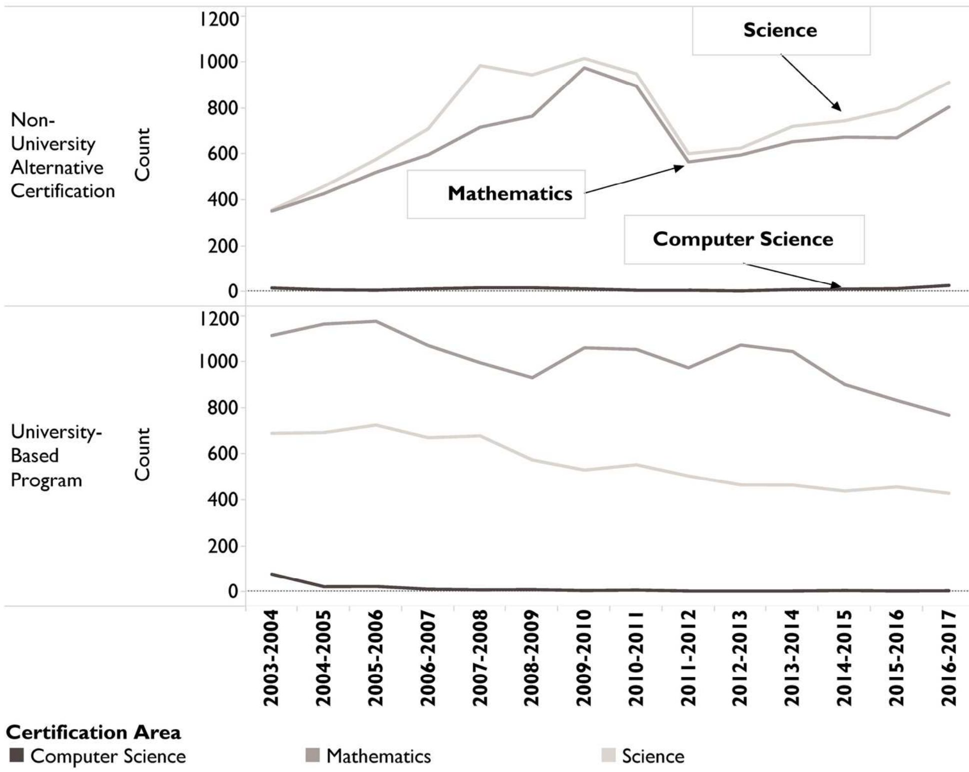

Figure 2: STEM teacher production in Texas from 2004 until 2017, comparing production from regular and alternative certification pathways. Data from Texas Education Agency.

152x120mm (300 x 300 DPI)

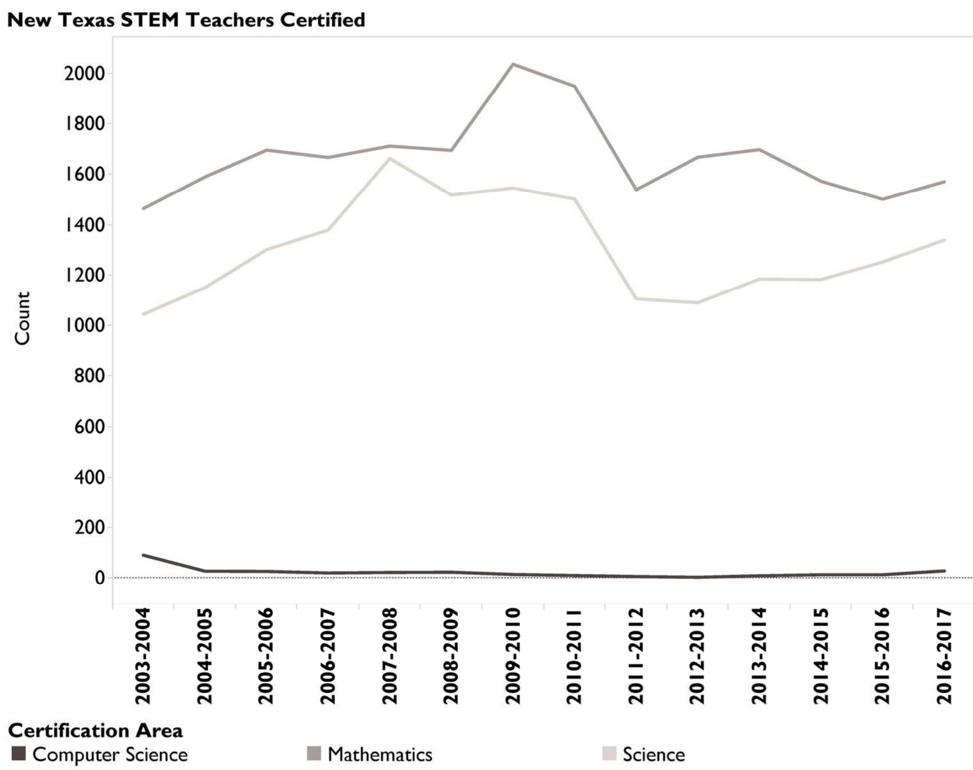

Figure 3: STEM teacher production in Texas from 2004 until 2017, summing all pathways. Data from Texas Education Agency.

151x118mm (300 x 300 DPI)

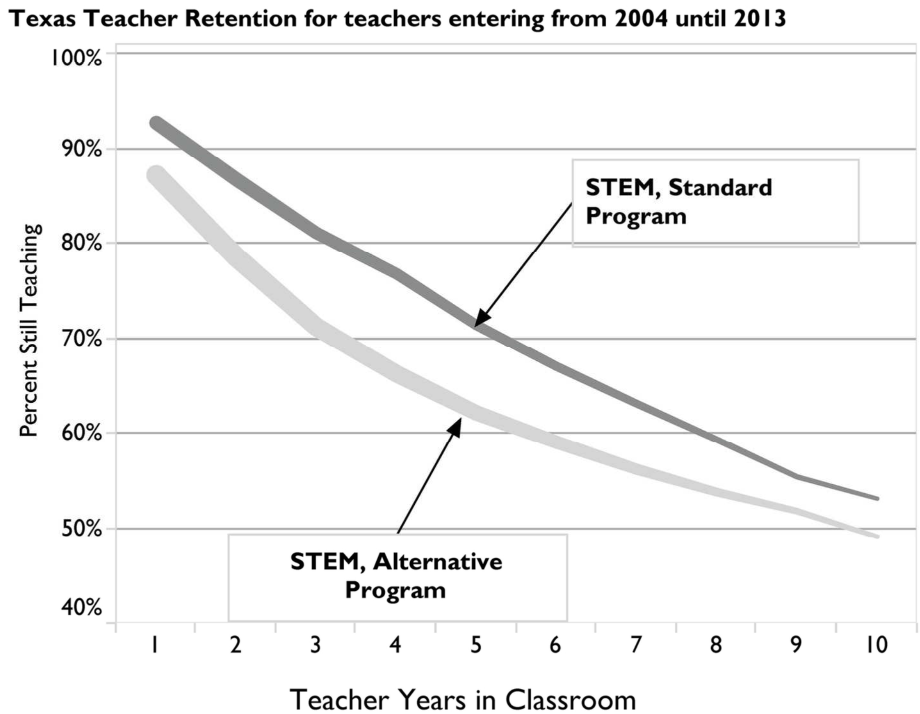

Figure 4: Retention of Texas STEM teachers in teaching by preparation pathway, averaged over cohorts entering from 2004 until 2013. Data from Texas Education Agency through Education Research Center.

142x107mm (300 x 300 DPI)

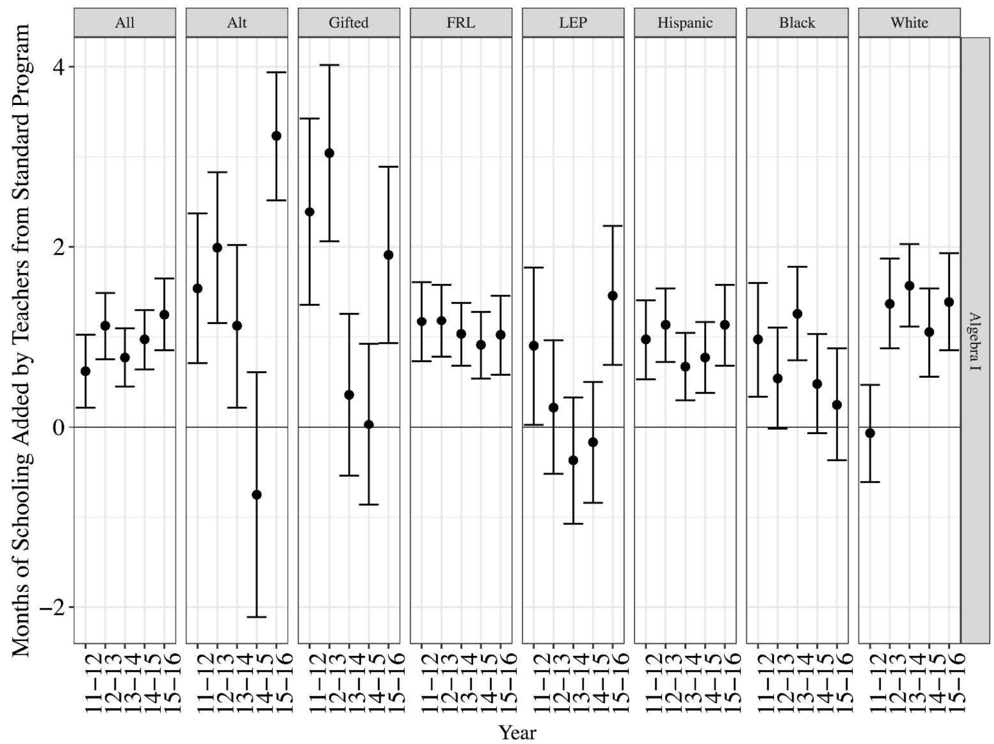

Figure 5: Added months of schooling for all Algebra I teachers from standard programs as compared with teachers from alternative programs, overall and for student subgroups, showing change over time. Data from Texas Education Agency through Education Research Center.

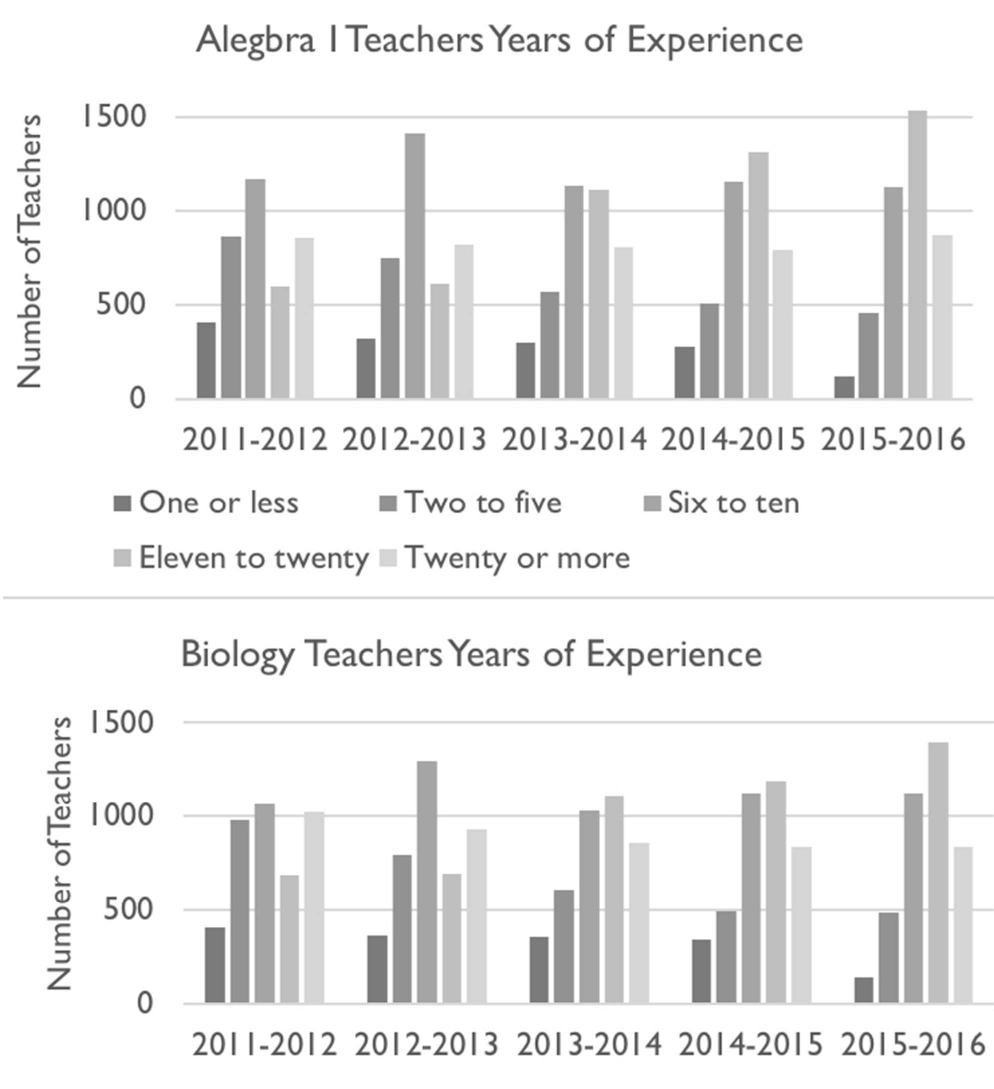

Figure 6: Distribution of teacher years of experience in Algebra I and Biology. Data from Texas Education Agency through Education Research Center.

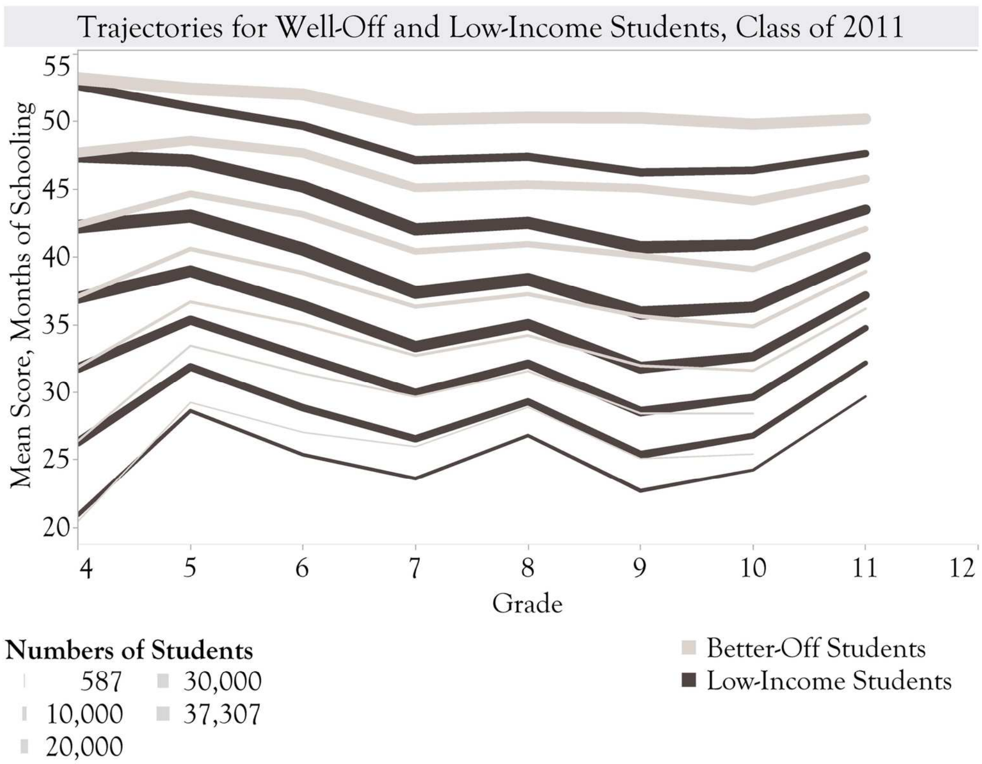

Figure 7: Mathematics exam averages in Months of Schooling units over time for cohorts of students from the Class of 2011 grouped both by their scores in 4th grade, and by their free/reduced lunch status. Data from Texas Education Agency through Education Research Center.

148x114mm (300 x 300 DPI)

Table 1: STEM teachers out of field in main assignment or not certified. Source: Schools and Staffing Survey (2012).

Subject	***Number of Teachers***	***Percent with no major in main assignment or not certified***
Mathematics	144,800	38%
Science	126,300	27%
Biology	51,900	35%
Physical Science	64,600	62%
Chemistry	24,300	66%
Earth Sciences	12,400	68%
Physics	13,300	63%

Table 2: Percentages of teachers who followed various pathways. Percentages in each block and each column sum to 100%. Rows shaded in grey are included in our Standard category, while all others are designated as Alternative.

Subject	Years of Experience	First Cert	Cert Program	Type	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016
Algebra I	<20	Standard	Standard	IHE	35%	37%	38%	39%	39%
Algebra I	<20	Probationary	Alternative	Not IHE	42%	43%	43%	43%	43%
Algebra I	<20	Probationary	Alternative	IHE	8%	7%	6%	7%	7%
Algebra I	<20	Standard	Alternative	Not IHE	1%	1%	1%	1%	1%
Algebra I	<20	Probationary	Post-Bacc	IHE	7%	6%	6%	6%	6%
Algebra I	<20	Standard	Post-Bacc	IHE	6%	5%	5%	4%	4%
Algebra I	<20	Standard	Alternative	IHE	1%	1%	1%	1%	1%
Algebra I	<4	Standard	Standard	IHE	16%	19%	20%	22%	26%
Algebra I	<4	Probationary	Alternative	Not IHE	64%	65%	64%	62%	59%
Algebra I	<4	Probationary	Alternative	IHE	6%	5%	3%	2%	3%
Algebra I	<4	Standard	Alternative	Not IHE	2%	2%	4%	4%	3%
Algebra I	<4	Probationary	Post-Bacc	IHE	5%	2%	2%	3%	4%
Algebra I	<4	Standard	Post-Bacc	IHE	5%	5%	5%	4%	4%
Algebra I	<4	Standard	Alternative	IHE	2%	1%	1%	1%	<1
Biology	<20	Standard	Standard	IHE	25%	25%	25%	24%	25%
Biology	<20	Probationary	Alternative	Not IHE	48%	49%	50%	51%	51%
Biology	<20	Probationary	Alternative	IHE	8%	8%	8%	8%	8%
Biology	<20	Standard	Alternative	Not IHE	1%	1%	1%	2%	2%
Biology	<20	Probationary	Post-Bacc	IHE	8%	8%	7%	7%	7%
Biology	<20	Standard	Post-Bacc	IHE	8%	8%	7%	7%	7%
Biology	<20	Standard	Alternative	IHE	1%	1%	1%	1%	1%
Biology	<4	Standard	Standard	IHE	10%	12%	13%	14%	16%
Biology	<4	Probationary	Alternative	Not IHE	66%	64%	65%	61%	64%
Biology	<4	Probationary	Alternative	IHE	7%	7%	7%	6%	6%
Biology	<4	Standard	Alternative	Not IHE	2%	3%	4%	4%	3%
Biology	<4	Probationary	Post-Bacc	IHE	6%	5%	4%	5%	4%
Biology	<4	Standard	Post-Bacc	IHE	8%	6%	6%	6%	5%
Biology	<4	Standard	Alternative	IHE	<1%	1%	1%	2%	1%

Table 3: Numbers of teachers in sample.

Subject	Years of Experience	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016
Algebra I	<20	2975	3055	3042	3180	3173
Algebra I	<4	814	657	544	502	364
Biology	<20	3079	3088	3040	3088	3086
Biology	<4	845	677	624	580	419

Table 4: Average characteristics of students and years of experience for Standard and Alternative teachers with up to 20 years of experience, and for all years in the study. All differences between columns labeled Std and Alt are significant with p<0.001. Final columns show difference between classroom demographics of teachers from Standard and Alternative pathways after controlling for school identity.

Category	Discipline	Std.	Alt.	Std - Alt, School Control
EcoDis	Algebra I	50.9%	58.1% (0.05%)	-0.20% (0.08%) **
Gifted	Algebra I	3.6%	3.7% (0.02%)	0.16% (0.03%) ***
SpecEd	Algebra I	4.2%	4.1% (0.02%)	-0.19% (0.03%) ***
LEP	Algebra I	5.8%	8.1% (0.03%)	0.15% (0.04%) ***
Asian	Algebra I	1.9%	1.7% (0.01%)	0.19% (0.02%) ***
Black	Algebra I	12.3%	15.7% (0.03%)	-0.22% (0.06%) ***
Hispanic	Algebra I	51.6%	56.7% (0.05%)	0.10% (0.07%)
White	Algebra I	31.9%	23.9% (0.04%)	-0.10% (0.07%)
Yrs4-10	Algebra I	45.9%	48.3% (0.05%)	-2.40% (0.07%) ***
Yrs11-20	Algebra I	43.6%	23.2% (0.04%)	18.13% (0.07%) ***
EcoDis	Biology	43.5%	50.8% (0.04%)	-0.09% (0.08%)
Gifted	Biology	11.1%	11.0% (0.03%)	0.10% (0.05%)
SpecEd	Biology	3.3%	3.2% (0.01%)	0.07% (0.03%) **
LEP	Biology	4.2%	6.0% (0.02%)	-0.02% (0.03%)
Asian	Biology	3.6%	3.8% (0.02%)	0.02% (0.03%)
Black	Biology	10.7%	14.1% (0.03%)	-0.07% (0.05%)
Hispanic	Biology	47.4%	52.2% (0.04%)	0.09% (0.07%)
White	Biology	35.8%	27.7% (0.04%)	-0.05% (0.06%)
Track	Biology	26.3%	27.8% (0.06%)	0.41% (0.07%) ***
Yrs4-10	Biology	40.0%	49.3% (0.04%)	-8.54% (0.07%) ***
Yrs11-20	Biology	48.0%	24.2% (0.04%)	20.27% (0.07%) ***

* |t|>1.96, ** |t|>2.33, *** |t|>3.09

Table 5: Model (1) coefficients for Algebra I. Numbers in parentheses are uncertainties. Different variants of the model include terms as shown.

Random Effects	***1a***			***1b***			***1c***			***1d***		
Campus SD				9.16			8.84			8.84		
Teacher SD	10.97			8.18			8.03			8.03		
Class SD	6.15			5.75			5.64			5.34		
Fixed Effects												
Cert	1.42	(0.27)	***	1.11	(0.22)	***	0.94	0.22	***	0.90	(0.22)	***
EcoDis	-2.80	(0.05)	***	-2.73	(0.05)	***	-2.62	0.05	***	-2.62	(0.04)	***
Gifted	8.54	(0.11)	***	8.58	(0.11)	***	7.98	(0.11)	***	7.98	(0.11)	***
SpecEd	-10.38	(0.10)	***	-10.49	(0.10)	***	-9.82	(0.11)	***	-9.82	(0.11)	***
LEP	-6.21	(0.08)	***	-6.21	(0.08)	***	-6.03	(0.09)	***	-6.02	(0.11)	***
Asian	4.95	(1.57)	**	4.91	(1.56)	**	4.55	(1.56)	**	4.55	(1.56)	**
Black	-2.37	(1.56)		-2.28	(1.56)		-2.23	(1.56)		-2.23	(1.56)	
Hispanic	-1.80	(1.55)		-1.74	(1.55)		-1.73	(1.59)		-1.73	(1.59)	
White	0.21	(1.56)		0.24	(1.56)		0.19	(1.56)		0.19	(1.56)	
AveEcoDis							-3.30	0.26	***	-3.30	0.26	***
AveGifted							14.53	0.55	***	14.53	0.55	***
AveSpecEd							-3.40	0.25	***	-3.40	0.25	***
AveLEP							-1.36	0.29	***	-1.36	0.29	***
AveAsian							9.79	0.89	***	9.79	0.89	***
AveBlack							-2.93	0.41	***	-2.93	0.41	***
AveHispanic							-1.73	0.33	***	-1.73	0.33	***
AveWhite							0	0		0	0	
Tracked							-	-		-	-	
Yrs: 4-10										0.05	(0.16)	
Yrs: 11-20										0.24	(0.21)	
$\lambda_1(2012)$	36.48	(2.98)	***	37.23	(2.97)	***	35.81	(2.98)	***	35.95	(2.98)	***
$\lambda_1(2013)$	99.16	(3.23)	***	98.38	(3.22)	***	96.56	(3.23)	***	96.70	(3.23)	***
$\lambda_1(2014)$	55.48	(2.97)	***	54.41	(2.97)	***	52.83	(2.97)	***	52.80	(2.97)	***
$\lambda_1(2015)$	31.14	(2.92)	***	30.00	(2.92)	***	28.75	(2.92)	***	28.60	(2.92)	***
$\lambda_1(2016)$	98.35	(3.16)	***	96.43	(3.16)	***	95.87	(3.16)	***	95.60	(3.16)	***
$\lambda_2(2012)$	7.75	(5.35)		5.36	(5.35)		7.43	(6.03)		7.24	(6.03)	
λ_2 (2013)	10.35	(6.62)		11.53	(6.62)	**	14.51	(6.62)	**	14.23	(6.62)	**
$\lambda_2(2014)$	152.82	(6.08)	***	154.18	(6.08)	***	156.35	(6.08)	***	156.37	(6.08)	***
$\lambda_2(2015)$	215.93	(5.82)	***	217.02	(5.82)	***	218.28	(5.82)	***	218.51	(5.82)	***
$\lambda_2(2016)$	81.14	(6.47)	***	83.62	(6.47)	***	83.53	(6.47)	***	83.99	(6.47)	***
$\lambda_3(2012)$	71.19	(3.67)	***	72.20	(3.67)	***	71.36	(3.67)	***	71.43	(3.67)	***
$\lambda_3(2013)$	21.60	(4.27)	***	20.96	(4.27)	***	19.07	(4.27)	***	19.23	(4.27)	***
$\lambda_3(2014)$	-78.28	(3.97)	***	-78.86	(3.97)	***	-80.24	(3.96)	***	-80.24	(3.96)	***
$\lambda_3(2015)$	-116.6	(3.69)	***	-116.8	(3.69)	***	-117.5	(3.69)	***	-117.7	(3.69)	***
$\lambda_3(2016)$	-32.13	(4.19)	***	-33.28	(4.19)	***	-33.34	(4.19)	***	-33.59	(4.19)	***

* |t|>1.96, ** |t|>2.33, *** |t|>3.09

Table 6: Model (5) coefficients for Biology. Numbers in parentheses are uncertainties. Different variants of the model include terms as shown.

Random Effects	***1a***			***1b***			***1c***			***1d***		
Campus SD				8.16			7.24			7.24		
Teacher SD	9.89			6.89			5.98			5.98		
Class SD	5.62			5.32			4.67			4.67		
Fixed Effects												
Cert	1.31	(0.27)	***	0.53	(0.21)	**	0.27	(0.19)		0.24	(0.19)	
EcoDis	-2.87	(0.04)	***	-2.78	(0.04)	***	-2.34	(0.04)	***	-2.34	(0.04)	***
Gifted	8.05	(0.06)	***	8.13	(0.06)	***	6.32	(0.06)	***	6.32	(0.06)	***
SpecEd	-9.16	(0.09)	***	-9.24	(0.09)	***	-7.97	(0.10)	***	-7.97	(0.10)	***
LEP	-7.85	(0.08)	***	-7.82	(0.08)	***	-7.35	(0.08)	***	-7.35	(0.08)	***
Asian	4.15	(1.29)	**	4.04	(1.29)	**	3.39	(0.27)	***	3.39	(0.27)	***
Black	-1.30	(1.29)		-1.30	(1.29)		-1.02	(0.26)	***	-1.02	(0.26)	***
Hispanic	-1.47	(1.29)		-1.45	(1.29)		-1.27	(0.26)	***	-1.27	(0.26)	***
White	1.15	(1.29)		1.11	(1.29)		1.11	(0.26)	***	1.11	(0.26)	***
AveEcoDis							-6.49	(0.21)	***	-6.49	(0.21)	***
AveGifted							9.69	(0.22)	***	9.69	(0.22)	***
AveSpecEd							-5.38	(0.23)	***	-5.38	(0.23)	***
AveLEP							-2.67	(0.25)	***	-2.67	(0.25)	***
AveAsian							6.19	(0.51)	***	6.19	(0.51)	***
AveBlack							-5.59	(0.33)	***	-5.59	(0.33)	***
AveHispanic							-3.89	(0.26)	***	-3.89	(0.26)	***
AveWhite							0	0		0	0	
Tracked							5.91	(0.04)	***	5.91	(0.04)	***
Yrs: 4-10										-0.30	0.12	**
Yrs: 11-20										0.06	0.16	
$\lambda_1(2012)$	56.76	(2.59)	***	57.20	(2.59)	***	62.99	(2.57)	***	63.04	(2.58)	***
$\lambda_1(2013)$	-33.17	(2.67)	***	-33.30	(2.67)	***	-21.52	(2.65)	***	-21.31	(2.65)	***
$\lambda_1(2014)$	31.33	(2.64)	***	30.21	(2.64)	***	41.73	(2.62)	***	41.70	(2.62)	***
$\lambda_1(2015)$	17.00	(2.62)	***	15.10	(2.62)	***	24.77	(2.60)	***	24.74	(2.61)	***
$\lambda_1(2016)$	23.19	(2.60)	***	20.38	(2.60)	***	30.60	(2.58)	***	30.48	(2.59)	***
$\lambda_2(2012)$	-160.9	(4.85)	***	-162.3	(4.85)	***	-163.9	(4.82)	***	-164.0	(4.82)	***
λ_2 (2013)	230.46	(5.01)	***	230.25	(5.01)	***	210.15	(4.98)	***	209.80	(4.98)	***
$\lambda_2(2014)$	97.84	(4.88)	***	99.38	(4.88)	***	79.49	(4.85)	***	79.54	(4.85)	***
$\lambda_2(2015)$	129.18	(4.77)	***	131.93	(4.77)	***	117.81	(4.74)	***	117.85	(4.74)	***
$\lambda_2(2016)$	146.59	(4.77)	***	150.87	(4.77)	***	135.70	(4.74)	***	135.87	(4.74)	***
$\lambda_3(2012)$	198.33	(2.79)	***	199.27	(2.79)	***	193.72	(2.77)	***	193.76	(2.77)	***
$\lambda_3(2013)$	-81.86	(2.99)	***	-81.65	(2.99)	***	-75.51	(2.97)	***	-75.33	(2.97)	***
$\lambda_3(2014)$	-19.13	(2.87)	***	-19.86	(2.86)	***	-13.76	(2.85)	***	-13.78	(2.85)	***
$\lambda_3(2015)$	-29.57	(2.74)	***	-30.87	(2.74)	***	-28.96	(2.72)	***	-28.99	(2.72)	***
$\lambda_3(2016)$	-53.73	(2.76)	***	-55.85	(2.76)	***	-53.75	(2.74)	***	-53.84	(2.74)	***

* |t|>1.96, ** |t|>2.33, *** |t|>3.09

Table 7: Months of added student learning per year in classes of standard certified teachers, broken down by discipline, school year, teacher experience, teacher preparation pathway, and using the three different multilevel models given in Equations (2)-(4).

Discipline	Year	Experience	Random			Fixed			None		
Algebra I	11-12	<20	0.62	(0.40)		0.38	(0.44)		1.05	(0.49)	*
Algebra I	12-13	<20	1.12	(0.37)	**	1.18	(0.41)	**	1.21	(0.47)	**
Algebra I	13-14	<20	0.77	(0.32)	**	0.58	(0.35)		1.08	(0.40)	**
Algebra I	14-15	<20	0.97	(0.33)	**	0.73	(0.36)	*	1.21	(0.41)	**
Algebra 1	15-16	<20	1.25	(0.40)	***	0.93	(0.44)	**	1.99	(0.51)	***
Biology	11-12	<20	0.15	(0.32)		0.29	(0.35)		-0.09	(0.41)	
Biology	12-13	<20	0.48	(0.36)		0.27	(0.40)		0.85	(0.45)	
Biology	13-14	<20	0.51	(0.28)		0.46	(0.30)		0.62	(0.35)	
Biology	14-15	<20	0.28	(0.29)		0.30	(0.32)		0.51	(0.38)	
Biology	15-16	<20	0.36	(0.30)		0.52	(0.32)		0.48	(0.36)	
Algebra I	11-12	<4	-1.18	(0.94)		-0.39	(1.24)		-1.67	(1.08)	
Algebra I	12-13	<4	0.17	(1.04)		2.15	(1.38)		-0.90	(1.21)	
Algebra I	13-14	<4	0.73	(1.03)		0.81	(1.43)		0.63	(1.19)	
Algebra I	14-15	<4	1.42	(0.99)		0.42	(1.36)		1.82	(1.17)	
Algebra I	15-16	<4	0.93	(1.67)		0.26	(3.18)		0.89	(1.80)	
Biology	11-12	<4	1.28	(0.78)		0.30	(0.98)		1.92	(0.95)	*
Biology	12-13	<4	-0.60	(0.95)		-1.20	(1.28)		-0.54	(1.08)	
Biology	13-14	<4	0.07	(0.85)		-0.20	(1.13)		-0.09	(0.96)	
Biology	14-15	<4	-0.42	(0.92)		0.82	(1.34)		-0.73	(1.00)	
Biology	15-16	<4	-0.21	(1.14)		-3.08	(2.19)		-0.02	(1.18)	

* |t|>1.96, ** |t|>2.33, *** |t|>3.09

Table 8: Months of Schooling gained by student subgroups in classrooms of standard certified teachers with up to 20 years of experience using model (5).

Year	**Subgroup**	**Algebra**			**Biology**		
11-12	Gifted	2.39	(1.03)	**	0.63	(0.64)	
11-12	Alt	1.54	(0.83)		0.93	(0.85)	
11-12	FRL	1.17	(0.44)	**	0.17	(0.36)	
11-12	LEP	0.90	(0.87)		0.36	(0.83)	
11-12	Black	0.97	(0.63)		0.70	(0.55)	
11-12	Hispanic	0.97	(0.44)	*	0.12	(0.37)	
11-12	White	-0.07	(0.54)		-0.28	(0.44)	
12-13	Gifted	3.04	(0.98)	***	-0.28	(0.60)	
12-13	Alt	1.99	(0.84)	**	-1.00	(0.83)	
12-13	FRL	1.18	(0.40)	**	-0.09	(0.35)	
12-13	LEP	0.22	(0.74)		0.93	(0.79)	
12-13	Black	0.54	(0.56)		0.55	(0.52)	
12-13	Hispanic	1.13	(0.41)	**	-0.18	(0.37)	
12-13	White	1.37	(0.50)	**	-0.02	(0.41)	
13-14	Gifted	0.36	(0.90)		1.10	(0.49)	*
13-14	Alt	1.12	(0.90)		0.29	(0.87)	
13-14	FRL	1.03	(0.35)	**	0.65	(0.32)	*
13-14	LEP	-0.37	(0.70)		0.10	(0.74)	
13-14	Black	1.26	(0.52)	**	0.12	(0.48)	
13-14	Hispanic	0.67	(0.37)		0.53	(0.32)	
13-14	White	1.57	(0.46)	***	0.25	(0.37)	
14-15	Gifted	0.03	(0.89)		0.66	(0.51)	
14-15	Alt	-0.75	(1.36)		1.18	(1.44)	
14-15	FRL	0.91	(0.37)	**	0.38	(0.34)	
14-15	LEP	-0.17	(0.67)		-0.15	(0.50)	
14-15	Black	0.48	(0.55)		-0.19	(0.55)	
14-15	Hispanic	0.77	(0.39)	**	0.40	(0.34)	
14-15	White	1.05	(0.49)	*	0.33	(0.41)	
15-16	Gifted	1.91	(0.98)	*	0.78	(0.49)	
15-16	Alt	3.23	(0.71)	***	1.73	(0.66)	**
15-16	FRL	1.02	(0.44)	**	0.45	(0.35)	
15-16	LEP	1.46	(0.77)		0.46	(0.70)	
15-16	Black	0.25	(0.62)		-0.24	(0.50)	
15-16	Hispanic	1.13	(0.45)	**	0.31	(0.33)	
15-16	White	1.39	(0.54)	**	0.25	(0.39)	

* |t|>1.96, ** |t|>2.33, *** |t|>3.09

Table 9: Months of Schooling gained by student subgroups in classrooms of standard certified teachers with less than 4 years of experience using model (5).

Year	**Subgroup**	**Algebra I**			**Biology**		
11-12	Gifted	2.26	(2.32)		1.20	(1.40)	
11-12	Alt	-2.88	(2.00)		4.13	(1.85)	*
11-12	FRL	-0.79	(0.99)		1.55	(0.82)	
11-12	LEP	-2.68	(1.79)		4.05	(1.79)	**
11-12	Black	0.36	(1.52)		2.36	(1.20)	*
11-12	Hispanic	-1.28	(1.01)		1.81	(0.86)	*
11-12	White	-2.48	(1.43)		0.28	(1.04)	
12-13	Gifted	2.25	(2.53)		-2.36	(1.61)	
12-13	Alt	-0.42	(2.28)		-1.47	(1.99)	
12-13	FRL	-0.13	(1.10)		-0.34	(0.99)	
12-13	LEP	0.05	(1.82)		0.54	(1.97)	
12-13	Black	-1.32	(1.47)		-1.17	(1.33)	
12-13	Hispanic	-0.06	(1.11)		-0.45	(1.03)	
12-13	White	0.72	(1.42)		0.14	(1.22)	
13-14	Gifted	0.27	(2.50)		1.15	(1.54)	
13-14	Alt	0.64	(2.48)		1.31	(2.38)	
13-14	FRL	1.57	(1.11)		0.03	(0.91)	
13-14	LEP	-0.25	(1.98)		-0.01	(1.88)	
13-14	Black	0.85	(1.46)		-1.99	(1.30)	
13-14	Hispanic	1.51	(1.14)		0.36	(0.93)	
13-14	White	2.16	(1.44)		-0.28	(1.14)	
14-15	Gifted	-2.09	(2.42)		1.35	(1.55)	
14-15	Alt	6.49	(3.32)	*	2.60	(3.05)	
14-15	FRL	1.34	(1.08)		-1.04	(1.04)	
14-15	LEP	-0.42	(1.80)		0.28	(1.85)	
14-15	Black	3.75	(1.45)	**	-1.56	(1.27)	
14-15	Hispanic	0.78	(1.11)		-0.15	(1.01)	
14-15	White	-0.47	(1.50)		0.26	(1.23)	
15-16	Gifted	1.09	(3.17)		0.42	(1.92)	
15-16	Alt	8.42	(2.45)	***	3.65	(2.27)	
15-16	FRL	0.12	(1.75)		-1.40	(1.29)	
15-16	LEP	0.86	(2.55)		-5.02	(2.25)	*
15-16	Black	1.78	(2.28)		-1.00	(1.75)	
15-16	Hispanic	0.35	(1.80)		-0.66	(1.27)	
15-16	White	2.60	(2.02)		0.42	(1.49)	

* $|t|>1.96$, ** $|t|>2.33$, *** $|t|>3.09$

Table 10: Probability of being reassigned to teach Algebra I or Biology in the same school given value-added score the previous year and certification pathway.

Year	**Discipline**	a_0			a_T			a_c		
2012-2013	Algebra I	-0.17	(0.05)	***	0.26	(0.03)	***	0.20	(0.06)	**
2013-2014	Algebra I	-0.06	(0.05)		0.26	(0.03)	***	0.16	(0.06)	*
2014-2015	Algebra I	0.05	(0.05)		0.26	(0.03)	***	0.18	(0.06)	**
2012-2013	Biology	0.21	(0.04)	***	0.35	(0.03)	***	0.23	(0.06)	***
2013-2014	Biology	0.23	(0.04)	***	0.33	(0.03)	***	0.30	(0.06)	***
2014-2015	Biology	0.35	(0.05)	***	0.27	(0.03)	***	0.17	(0.06)	**

* p<.05, ** p<.01, *** p<.001