Teachers Without Borders: Consequences of Teacher Labor Force Mobility

Kevin C. Bastian
University of North Carolina at Chapel Hill
Gary T. Henry
Vanderbilt University

Many states have responded to teacher shortages by granting certification to individuals tradition- ally prepared out-of-state; now, out-of-state prepared teachers comprise a sizable percentage of the teacher workforce in many states. We know little about these teachers, and therefore, in the present study, we estimate the effectiveness of out-of-state prepared teachers in North Carolina elementary schools. We find that out-of-state prepared teachers are significantly less effective than in-state prepared and alternative entry teachers; however, there is a substantial overlap in the distributions of effectiveness across groups. Upon testing hypotheses to explain these findings, results indicate that differences in human capital help explain out-of-state prepared teachers' underperformance and suggest the utility of research evidence to inform state policy and local hiring decisions.

Keywords: *teacher labor markets, teacher preparation, teacher effectiveness*

Introduction

IN MANY states, population growth, teacher attri- tion, teacher retirements, and more employment opportunities for women have contributed to teacher shortages (Bacolod, 2007; Ingersoll & Smith, 2003). These shortages, coupled with a need for more highly effective teachers, have pushed many states to experiment with alterna- tives to solely licensing instructors prepared at in- state traditional education programs. Lateral/ alternative entry programs, which reduce barriers to employment by allowing individuals without teacher education credentials to complete require- ments for certification while concurrently teach- ing, have been a common state policy response (Feistritzer, 2011; National Research Council, 2010). For example, from 2000–2001 to 2009– 2010, the number of alternative entry teachers in North Carolina public schools increased 125%, from 6,626 to 15,028 (authors' analysis).

Through licensing agreements with national teacher accreditation and certification associa- tions, such as the National Council for the Accreditation of Teacher Education or the National Association of State Directors of Teacher Education and Certification, a frequently used but little studied alternative approach has been for states to grant more reciprocal teacher certification licenses and expand the number of out-of-state prepared teachers. This policy broad- ens the pool of potential teachers by facilitating interstate teacher mobility, and currently, many states rely on these imported teachers to staff a significant percentage of their classrooms. For example, in 2008–2009, 15 states reported that more than one third of the teachers to whom they granted initial certification were trained in another state (U.S. Department of Education, 2010). In North Carolina, more than 40% of ini- tially licensed teachers were trained out-of-state

TABLE 1
Individual and Workplace Characteristics

	Individual teacher characteristics		
Variable	Out-of-state prepared	In-state prepared	Alternative entry
% female	85.99	92.98**	85.59
	(34.71)	(25.54)	(35.14)
% minority	8.62	13.16**	33.68**
	(28.07)	(33.81)	(47.28)
Age	29.02	29.04	33.69**
	(7.11)	(7.89)	(8.82)
Teaching experience	1.93	1.80**	1.99
	(1.36)	(1.40)	(1.37)
Standardized exam scores	0.254	0.277*	0.049**
	(0.608)	(0.558)	(0.744)
Workplace characteristics			
Students' prior EOG scores	−0.091	−0.085	−0.344**
	(0.511)	(0.461)	(0.581)
School performance composite	60.34	60.47	54.58**
	(16.35)	(15.24)	(16.45)
School free- and reduced-price lunch %	54.72	57.96**	65.68**
	(25.82)	(23.47)	(25.34)
School minority %	53.33	45.81**	60.44**
	(26.29)	(27.55)	(28.58)
Short-term suspension rate (per 100 students)	8.11	8.06	12.00**
	(11.49)	(11.49)	(18.50)
Total per-pupil expenditures	8,475.41	8,512.11	8,828.33**
	(1,515.63)	(1,431.54)	(1,634.98)

Note. In Table 1, percentage female and minority identify unique teachers. Age, teaching experience, and standardized licensure exam scores identify unique teacher-year combinations. Students' prior EOG scores identify unique teacher-classroom combinations and school performance composite (number of tests passed divided by the number of tests taken), school free- and reduced-price lunch percentage, school minority percentage, short-term suspension rates, and total per-pupil expenditures identify unique teacher-school-year combinations. We tested the values for in-state and alternative entry teachers against those of out-of-state prepared teachers. EOG = end-of-grade.
*Statistical significance at the .05 level. **Statistical significance at the .01 level.

(U.S. Department of Education, 2010), and over the past decade, the number of out-of-state prepared instructors has increased 36%, from 21,316 to 29,006 (authors' analysis). (See Online Appendix Table 1, available at http://epa.sagepub.com/supplemental, for rates of reliance on out-of-state prepared teachers.)

Despite these three options (in-state prepared, out-of-state prepared, and alternative entry) for staffing schools, prior research on teacher effectiveness has generally combined in-state and out-of-state traditionally prepared teachers into a single category and compared the effectiveness of traditionally prepared teachers with that of alternatively prepared teachers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clotfelter, Ladd, & Vigdor, 2007, 2010; Kane, Rockoff, & Staiger, 2008). From a policy perspective, classifying in-state and out-of-state traditionally prepared teachers together ignores differential financial costs associated with these groups of instructors, including those of preparation, recruitment, development, and replacement, which make distinctions between the categories important (Alliance for Excellent Education, 2004). Furthermore, combining those traditionally prepared in-state with out-of-state

prepared teachers ignores potential state-to-state differences in the quality of teacher preparation programs and how labor markets work to sort teachers of lower quality out of exporting states (National Research Council, 2010).

Both theoretical and prior research evidence suggests three hypotheses predicting that out-of-state prepared teachers may be less effective than their in-state prepared peers: (a) differences in the educational environment, including curricula and standards, between the exporting and importing states (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009); (b) teacher candidates with lower levels of human capital—defined as cognitive ability and job-specific knowledge and skills—may need to be more mobile to find employment (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012); and (c) out-of-state prepared teachers who acquire human capital through on-the-job experience may become more competitive for positions back in their state of origin, causing high rates of teacher turnover and the potential for differential attrition (Boyd et al., 2006).

In this article, we examine the effectiveness of early-career out-of-state prepared teachers and address each hypothesis by answering the following questions:

Research Question 1: How does the effectiveness of traditionally prepared out-of-state teachers compare with that of individuals traditionally prepared in-state, or admitted into the profession through alternative entry programs?

Research Question 2: If effectiveness differences exist between out-of-state prepared teachers and the groups specified above, what accounts for those performance disparities—less familiarity with the educational environment, lower levels of human capital, and/or the attrition patterns of out-of-state prepared teachers?

By way of preview, we find that, on average, out-of-state prepared teachers consistently underperform in-state traditionally prepared and alternative entry teachers in elementary grades mathematics and reading. There is considerable heterogeneity in the performance of out-of-state prepared teachers, however, and a substantial overlap in the distributions of effectiveness across our teacher groups. Upon testing hypotheses to better understand these results, we find that out-of-state prepared teachers are initially less effective than their in-state prepared peers and do not narrow this effectiveness gap during their early-career period. Furthermore, results show heterogeneity in teacher effectiveness across states—states with greater competition for open teaching positions export less effective teachers. Taken together, these results indicate that differences in human capital help explain the underperformance of out-of-state prepared teachers. Last, we find that out-of-state prepared teachers exit North Carolina public schools at significantly higher rates than their in-state prepared peers and that those who exit are significantly less effective than those who stay. Overall, these analyses highlight the importance of teacher labor markets to student achievement and suggest the utility of evidence on teacher effectiveness and attrition to inform state policy and local hiring decisions.

In the following sections, we first summarize the research evidence on the effects of teacher preparation and then present the data and sample, methods, and results for our first research question. Next, for each of our three hypotheses, we detail the theoretical and empirical evidence for the hypothesis, our methods to test the hypothesis, and results. Finally, we conclude with a discussion of potential policy responses.

Prior Research on the Effects of Teacher Preparation

Within the past decade, a number of studies have compared the effectiveness of traditionally and alternatively certified/prepared teachers, and overall, this body of research has generated two broad findings. First, teachers holding regular certification/traditional preparation appear to be more effective in the early stages of their careers (Boyd et al., 2006; Clotfelter et al., 2007, 2010; Goldhaber & Brewer, 2000; Henry et al., in press; Kane et al., 2008). These returns to certification/preparation fade quickly, however, such that the efficacy of the credential as a signal of teacher quality is limited (Boyd et al., 2006). Second, there is more variation in teacher

effectiveness within certification/preparation categories than between them, meaning factors outside preparation may better determine a teacher's classroom success (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Henry et al., in press; Kane et al., 2008). Despite these findings, four issues provide justification for further research on the relative effectiveness of teachers prepared through different programs or entering through different teacher preparation routes (e.g., in-state, out-of-state, and alternative entry).

First, there is substantial heterogeneity in preparation components and requirements within traditional and alternative routes, yet many studies only include these two broad categories in their analyses (Darling-Hammond, Berry, & Thoreson, 2001; Greenberg, Pomerance, & Walsh, 2011; National Research Council, 2010). Second, to best measure the influence of teacher preparation, research should focus on teachers early in their careers, when the effects of preparation are the strongest (Goldhaber, Liddle, & Theobald, 2013; Henry et al., in press). To date, however, only a few studies limit their analyses to this early-career sub-sample (Boyd et al., 2006; Henry et al., in press; Kane et al., 2008), whereas others draw inferences about the value of teacher preparation based on the effectiveness of teachers more than a decade removed from formal training (Clotfelter et al., 2007, 2010; Goldhaber & Brewer, 2000). Third, due to teacher selection into training, great care must be taken when interpreting preparation estimates. Without a measure of teachers' academic ability prior to entering a teacher preparation category, estimates of teacher preparation effectiveness combine the effects of selection and preparation (Boyd et al., 2006).

Finally, when examining the effects of teacher preparation, researchers must consider how labor market factors influence estimates. Examples of such factors include (a) teachers' preferences to secure employment close to home or their preparation university (Boyd et al., 2005; Reininger, 2012), (b) teachers' preferences to work in schools with high-quality working conditions and schools' preferences to hire better credentialed teachers (Boyd, Lankford, Loeb, Ronfeldt, & Wyckoff, 2011; Hanushek, Kain, & Rivkin, 2004), and (c) the non-random attrition of teachers from the profession (Goldhaber, Gross, & Player, 2011; Henry, Bastian, & Fortner, 2011).

To better understand why out-of-state prepared teachers underperform, we will discuss each of these labor market factors in greater detail in our hypothesis tests section.

Although research on traditionally and alternatively certified/prepared teachers is expanding, only two studies separately assess the effectiveness of out-of-state prepared teachers: Henry and colleagues (in press) in North Carolina, and Goldhaber and colleagues (2013) in Washington State. Henry and colleagues found that teachers prepared out-of-state were significantly less effective than traditionally prepared in-state teachers in elementary grades mathematics and reading, where out-of-state prepared teachers are highly concentrated, and in high school mathematics and science. Although this work reported effect estimates for out-of-state prepared teachers compared with in-state traditionally prepared teachers, the study did not (a) compare the effectiveness of out-of-state prepared teachers with that of teachers from any other route into the profession, such as alternative entry teachers or (b) investigate explanations for these performance disparities. In Washington, Goldhaber and colleagues found that some in-state teacher preparation programs outperformed out-of-state prepared teachers in reading, but overall, there were few effectiveness differences based on teachers' prior preparation (Goldhaber et al., 2013).

Building upon these previous studies, in the next section, we present the data and sample, methods, and results concerning the comparative effectiveness of out-of-state prepared teachers in North Carolina elementary schools.

Research Question 1: How Effective Are Out-of-State Prepared Teachers?

Data and Sample

Our first objective in the present study was to estimate the comparative effectiveness of out-of-state prepared teachers in elementary school mathematics and reading. This required developing a teacher preparation coding scheme, matching students to their classroom teachers; building longitudinal analysis files with student, teacher, classroom, and school characteristics; and estimating effects. Below, we detail our classification of teachers, study sample, and the covariates used in analyses to address our first research question.

166

Classification of Teachers. For this analysis, we classified teachers into exclusive and fixed categories according to their formal preparation—earning a degree or completion of a certificate program—most proximate to entering the profession. Specifically, we coded three teacher preparation categories as follows: (a) out-of-state prepared (fully certified upon first entering the teaching profession after earning an undergraduate/graduate degree or completing a licensure/certificate program at an out-of-state university), (b) in-state prepared (fully certified upon first entering the teaching profession after earning an undergraduate/graduate degree or completing a licensure/certificate program at an in-state university), and (c) alternative entry (not fully certified upon first entering the teaching profession).[1] To classify teachers into these groups, we used certified salary files from the North Carolina Department of Public Instruction (NCDPI) to identify the year an individual began teaching. Next, using the NCDPI licensure audit file, we determined the basis for a teacher's original teaching license. If this initial basis indicated that a teacher had not completed all licensure requirements upon first entering the profession, we classified her as alternative entry. Finally, using education files from the University of North Carolina General Admission and the NCDPI, we determined an individual's graduation year, degree level, and degree origin and assigned individuals to the single teacher preparation category most proximate to their entry into the profession.

One limitation of this coding scheme is that our categories are entirely based on post-secondary (teacher preparation) data. This has implications for our overall and hypothesis test results because within our three groups, there may be important distinctions between those teachers who (a) never had any schooling (K–12 or post-secondary) in North Carolina, (b) had K–12 schooling outside North Carolina but were prepared to teach in-state, (c) had K–12 schooling in North Carolina but were prepared to teach out-of-state, and (d) had all their schooling in North Carolina. Unfortunately, data are not available for us to track the K–12 schooling of teachers in our sample. As more states create PK–20 data systems, however, such analyses will be valuable.

Study Sample. The analysis data for our first research question span the 2005–2006 through 2009–2010 school years and are limited to teachers with less than 5 years teaching experience. We restrict our sample to these early-career teachers for three reasons. First, teachers are most likely to display measurable and relevant preparation influences early in their careers, before on-the-job learning dilutes the effects of preparation. Second, for states considering policy mechanisms to improve the quality of their teacher workforce, individuals who recently entered the profession entail particular significance because they comprise more than 25% of the teaching population (Ingersoll & Merrill, 2010). Finally, we limit our sample because out-of-state prepared teachers are highly concentrated in North Carolina elementary schools, representing nearly 37% of the state's early-career, tested-grades (3–5) instructors.

The key data feature for this analysis is our use of actual classroom rosters, which allowed us to match students to approximately 93% of individual instructors over the 5-year study period, construct classroom level covariates, and account for multiple teachers within a subject-year for a given student. Table 1 presents individual and workplace characteristics for the teachers in our value-added analyses. Regarding individual teacher characteristics, Table 1 indicates that in comparison with out-of-state prepared teachers (a) in-state prepared teachers are more likely to be female and a racial/ethnic minority, have slightly less teaching experience, and slightly higher standardized exam scores; and (b) alternative entry teachers are more likely to be a racial/ethnic minority, are older, and have lower standardized exam scores. Data for workplace characteristics show that in comparison with out-of-state prepared teachers: (a) in-state prepared teachers work in schools with more students qualifying for subsidized school lunches but fewer minority students, and (b) alternative entry teachers work in schools that are lower performing, have higher suspension rates, and have more students that are a racial/ethnic minority and qualify for subsidized school lunches. Alternative entry teachers also teach students with significantly lower prior exam scores.

Dependent Variables and Covariates. For this analysis, we based students' prior and current test

score performance on the North Carolina Grade 3 pre-test and the End-of-Grade (EOG) mathematics and reading exams in Grades 3 to 5. These EOG exams are criterion-referenced, vertically equated across years to allow for meaningful comparisons, and have been rigorously analyzed to ensure valid psychometric properties (NCDPI, 2011). To remove any secular trends, we standardized all tests within subject, grade, and year and included year fixed effects in our model specifications. We used indicators for in-state prepared and alternative entry teachers, in comparison with out-of-state prepared teachers (reference group), to answer our first research question. Finally, to adjust for factors influencing teacher effectiveness outside the control of these teacher preparation categories, we included a rich set of student, classroom, teacher, and school covariates (shown in Table 2) in our models.

Method

Adjusted Average Teacher Value-Added. To estimate the adjusted-average effectiveness differences between early-career out-of-state prepared teachers and in-state prepared and alternative entry teachers, we utilize our extensive administrative database to specify an ordinary least squares (OLS) value-added model with a rich set of student, classroom, teacher, and school covariates. We use cluster-adjusted standard errors at the school-year level to account for the nesting of students and teachers within schools that, if left unadjusted, could result in reduced standard errors and significance tests that produce false positives. Overall, we prefer this model and its rich set of controls because (a) it allows us to make teacher effectiveness comparisons across our entire (statewide) sample (we discuss specification checks below); (b) it does not require multiple years of prior student test records, which would eliminate grade(s) from our analysis; and (c) it does not require multiple years of teacher effectiveness data, which would limit our focus on early-career teachers. We specified the following equation to estimate adjusted-average teacher preparation category effects:

$$Y_{ijst} = \beta_0 + \beta_1 Y_{it-n} + \beta_2 Instate + \beta_3 Alternative$$
$$+ \beta_x X_{ijst} + \beta_z Z_{jst} + \beta_w W_{st} + \varepsilon_{ijst}, \quad (1)$$

where Y_{ijst} is the test score for student i in classroom j in school s at time t; β_2 and β_3 estimate the average effect of in-state prepared and alternative entry instructors, respectively, relative to out-of-state prepared teachers; *In-state* and *Alternative* are indicator variables that equal 1 if the teacher entered teaching through that category and 0 if not; Y_{it-n} represents the mathematics and reading scores for student i from the prior grade[2]; X_{ijst} represents a set of individual student covariates; Z_{jst} represents a set of classroom and teacher covariates; W_{st} represents a set of school covariates; and ε_{ijst} is a disturbance term representing all unexplained variation.

In response to our preferred rich-covariate adjustment model, the fundamental question is how well it controls for endogeneity threats. We prefer the OLS approach because rich-covariate adjustment models estimate teacher preparation category effects based on the entire (statewide) sample of teachers, rather than the more limited sample of within-unit variation in fixed effects approaches. However, if unmeasured school factors or the non-random assignment of teachers to students affect student test performance and are both correlated with the preparation categories and omitted from models, then fixed effects will produce preferred, internally valid estimates.

Therefore, to assess the robustness of our preferred OLS value-added model, we employ three fixed effects specifications—school fixed effects, school-by-year fixed effects, and student (levels model) fixed effects. Here, the school fixed effects limit comparisons to teachers within the same school, thereby eliminating any uncontrolled, time-invariant school factors that may influence estimates, whereas the school-by-year fixed effects restrict teacher preparation comparisons to teachers in the same school and year to control for unmeasured school and temporal trends. Our student fixed effects model is a levels (no prior test score) specification that uses students as their own control and compares students' test score outcomes (deviation from the students' mean scores standardized by grade and year) when taught by an out-of-state prepared teacher to outcomes when instructed by an in-state prepared or alternative entry teacher. These fixed effects models continue to include a rich set of student, classroom, teacher, and school

168

TABLE 2
Covariates for Analyses

Student covariates	Classroom and teacher covariates	School covariates
1. Prior test scores (reading and math)	16. Class size	21. School size
2. Classmates prior test scores	17. Heterogeneity of prior achievement within the classroom	22. School size squared
3. Days absent	18. Teacher out-of-field status	23. Violent acts per 1,000 students
4. Structural mobility	19. Single-year indicators for teacher experience	24. Short-term suspension rate
5. Within year mobility	20. Teacher preparation categories	25. Total per-pupil expenditures
6. Other between year mobility	a. Out of state prepared (reference group)	26. District teacher supplements
7. Race/ethnicity	b. In-state prepared	27. Racial/ethnic composition
8. Gender	c. Alternative entry	28. Concentration of poverty
9. Poverty status		
10. Gifted		
11. Disability		
12. Currently limited English proficient		
13. Was limited English proficient		
14. Overage for grade		
15. Underage for grade		

Note. We included these covariates in our preferred rich-covariate adjustment and school fixed effects analyses. Student fixed effects exclude time-invariant student characteristics; school-year fixed effects exclude school-level variables.

covariates to isolate the effect of preparation categories—we exclude time-invariant student characteristics from the student fixed effects models and school covariates from the school-by-year fixed effects models. Because these fixed effects approaches only identify coefficients for teacher preparation categories based on within-unit (school, school-by-year, or student) variation, the results (Table 3) for our first research question include counts of unique teachers contributing to the estimates. Furthermore, all results tables provide observation counts for the student test records that actually contributed to the teacher preparation estimates—for instance, observation counts for school fixed effects models exclude schools in which during our study period no out-of-state prepared teachers worked or only out-of-state prepared teachers worked.

Individual Teacher-by-Year Value-Added. In addition to estimating the fixed, adjusted-average effectiveness differences between out-of-state prepared and in-state prepared and alternative entry instructors, we also examined the distribution of individual teacher-year effectiveness within and between these preparation categories. For this analysis, we estimated individual teacher-by-year effectiveness using an OLS value-added model with a rich set of student, classroom, and school covariates; we intentionally excluded teacher characteristics from these models to generate individual teacher-by-year estimates unadjusted for teacher covariates. Here, for each teacher-year, we specified the average of their students' residuals as the measure of individual effectiveness and aggregated these teacher-by-year estimates to display effectiveness distributions for each preparation category. Although we consider the adjusted-average effect estimates the most policy relevant comparison, these distributions of teacher effectiveness also suggest key policy implications.

Results

For our preferred rich-covariate adjustment model, Table 3 shows that both in-state prepared and alternative entry teachers significantly outperform out-of-state prepared teachers. In-state prepared instructors are

169

TABLE 3

Elementary School Mathematics and Reading Outcomes

Teacher preparation category	Rich-covariate adjustment		School fixed effects		School-year fixed effects		Student fixed effects	
	ES math	ES reading	ES math	ES reading	ES math	ES reading	ES math	ES reading
In-state prepared	0.024**	0.012**	0.019**	0.011**	0.019**	0.009**	0.023**	0.008*
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.005)	(0.004)
Alternative entry	0.030**	0.013*	0.029**	0.025**	0.021**	0.025**	0.019*	0.017**
	(0.007)	(0.005)	(0.007)	(0.005)	(0.008)	(0.006)	(0.009)	(0.006)
Observations used	715,172	1,008,362	638,290	883,837	492,354	673,522	120,677	170,691
Student covariates	✓	✓	✓	✓	✓	✓	✓	✓
Classroom/ teacher covariates	✓	✓	✓	✓	✓	✓	✓	✓
School covariates	✓	✓	✓	✓			✓	✓
Unique teacher counts contributing to preparation category estimates								
Out-of-state prepared	4,197	4,255	4,130	4,195	3,820	3,877	3,544	3,620
In-state prepared	5,330	5,413	4,509	4,603	3,672	3,753	3,577	3,682
Alternative entry	1,278	1,323	1,083	1,117	871	903	753	796

Note. All coefficients are in relation to out-of-state prepared teachers. ES = elementary school.
*Significance at the .05 level. **Significance at the .01 level.

0.024 and 0.012 standard deviations more effective in mathematics and reading, respectively; alternative entry teachers are 0.030 and 0.013 standard deviations more effective.[3] Using school, school-by-year, or student fixed effects to adjust for the sorting of students and teachers into schools or other unmeasured school characteristics does not alter the substance of these results. Although these significant results are not large, approximately 1% to 3% of a standard deviation, this effect size is comparable with findings from other early-career teacher preparation research (Boyd et al., 2006; Kane et al., 2008). To make these estimates more meaningful, we used the average gains in EOG scale score points between elementary grades and the average standard deviation on elementary grades EOG tests to convert the effects in Table 3 into equivalent days of schooling in a 180 day school year. Here, in comparison with students instructed by out-of-state prepared teachers, the in-state prepared effects of 0.024 and 0.012 add approximately 6 and 3.5 days of schooling in mathematics and reading, respectively; the alternative entry effects of 0.030 and 0.013 add approximately 7.5 and 3.7 days of schooling in mathematics and reading, respectively (Henry, Thompson, Fortner, Bastian, & Marcus, 2011). Overall, early-career out-of-state prepared teachers taught nearly 200,000 students in Grades 3 to 5 during our study period, which means that these results have practical significance for the achievement of the state's elementary grades students (see Online Appendix Table 2, available at http://epa.sagepub.com/supplemental, for all the coefficients from our rich-covariate adjustment models).
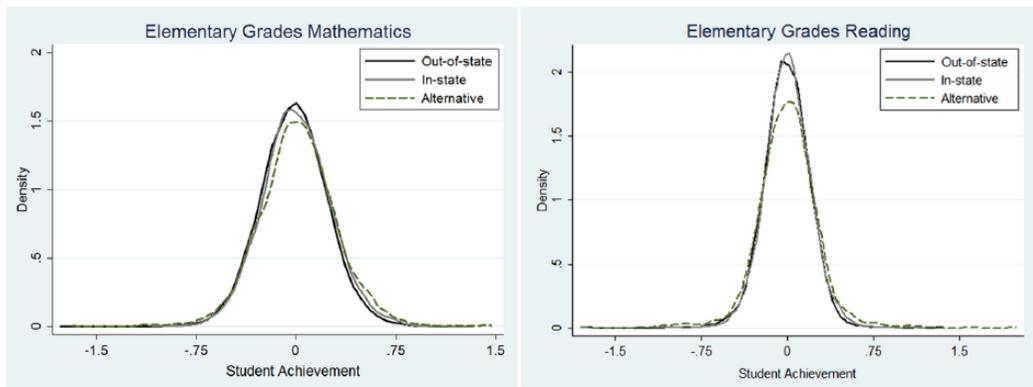
170

FIGURE 1. *Elementary grades teacher effectiveness distributions.*
Note. This figure displays individual teacher-by-year effectiveness for in-state prepared, out-of-state prepared, and alternative entry teachers.

To further contextualize the adjusted-average effectiveness differences between out-of-state prepared and in-state prepared and alternative entry teachers, Figure 1 displays the distribution of individual teacher-by-year effectiveness estimates for our three teacher preparation categories. Consistent with prior research, these figures illustrate that there is substantial overlap in the distribution of teacher effectiveness across groups (Boyd et al., 2007; Kane et al., 2008). In comparison with out-of-state prepared teachers, however, alternative entry instructors have a wider dispersion of effectiveness and both in-state prepared and alternative entry teachers have an effectiveness distribution shifted to the right. These figures complement our adjusted-average value-added results, showing that although out-of-state prepared teachers are less effective, on average, there are both many higher performing out-of-state prepared instructors and many lower performing in-state prepared, out-of-state prepared, and alternative entry teachers.

Research Question 2: Why Do Out-of-State Prepared Teachers Underperform?

To better understand why out-of-state prepared teachers are less effective, in the sections below, we present theoretical and research evidence, empirical tests, and results for each of our three hypotheses.

Hypothesis 1: Teachers' Lack of Familiarity With the Importing State

Background. As a result of the state standards/accountability movement and No Child Left Behind, each state has unique standards, curricula, and assessments on which their educational systems are built.[4] Importantly, the differences between states in these standards and assessments are substantial, and to the extent that teacher preparation programs are regulated by states, in-state colleges and schools of education may structure their course content and academic requirements to enable their graduates to become more familiar with state curricula and academic content than individuals prepared elsewhere (Carnoy & Loeb, 2002; Porter, Polikoff, & Smithson, 2009). In addition, through student teaching and other field experiences, pre-service teachers at in-state colleges and schools of education will have more opportunities to engage in teaching practice in educational environments—types of students, curricula/content, and schools—similar to their future, in-service classroom placements. Based on recent research which suggests that early-career instructors benefit from greater pre-service exposure to the school environments in which they will teach and the academic content they are expected to teach, we hypothesize that out-of-state prepared teachers will be less familiar with North Carolina's educational environment, and therefore, less effective than in-state prepared teachers (Boyd et al., 2009).
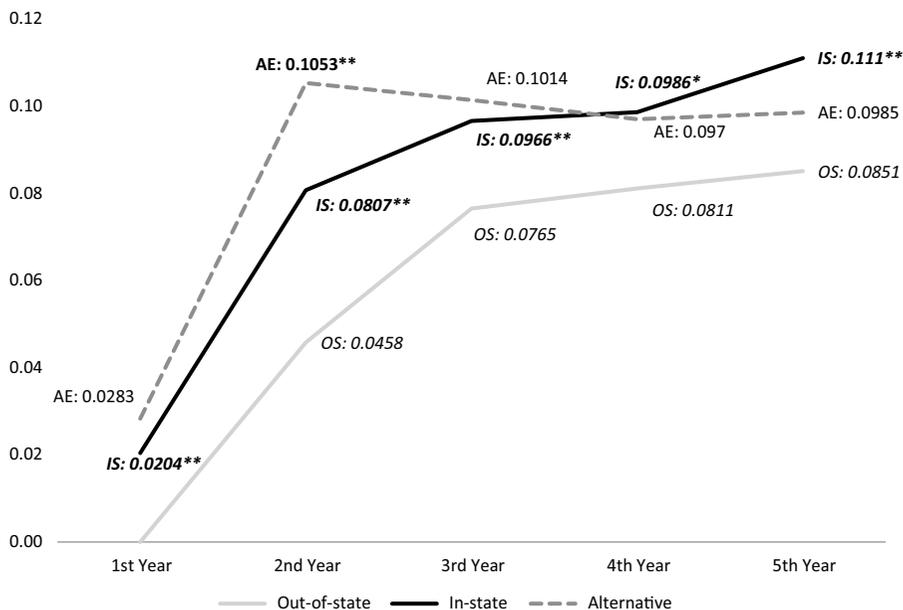
171

FIGURE 2. *Elementary grades math teacher effectiveness trajectories.*
Note. This figure displays teacher effectiveness relative to first-year out-of-state prepared teachers (coefficients shown in Table 4). For each year of teaching, we used joint significance tests to determine whether there are statistically significant differences between out-of-state prepared and in-state prepared and alternative entry teachers. IS = in-state; AE = alternative entry; OS = out-of-state.
*Statistically significant differences at the *p* < .05 level. **Significant differences at the .01 level.

Empirical Tests and Results. We test our lack of familiarity hypothesis in two main ways. First, based on the theory that lack of familiarity effects would be strongest in the first year of teaching and then lessen over time, as out-of-state prepared teachers become more familiar with the importing state, we re-ran the rich-covariate adjustment model shown in Equation 1 and included interactions between our teacher preparation categories and teacher experience indicators. Here, if a lack of familiarity is adversely influencing out-of-state prepared teacher performance, we would expect to see initial differences in effectiveness narrow over time. Instead, Figures 2 and 3 show (coefficients and *p* values reported in Table 4) that initial differences in effectiveness between out-of-state prepared and in-state prepared teachers widen in both mathematics and reading (particularly). The trajectories for alternative entry teachers are less stable, showing both rapid increases and decreases in performance. Overall, these graphs indicate that on-the-job exposure to a state's educational environment does not reduce initial effectiveness differences between in-state and out-of-state prepared teachers.

Second, using the Fordham Institute's ratings of states' mathematics and English language arts content standards, we tested whether out-of-state prepared teachers are less effective due to differences in content standards across states. Specifically, the Fordham Institute rates each state on the content and rigor of its standards using a 1 to 7 scale (Carmichael, Martino, Porter-Magee, & Wilson, 2010). We took these content ratings for mathematics and English language arts and created two variables: (a) the rating difference between a given state and North Carolina and (b) the absolute value of this rating difference. Next, we limited our sample to out-of-state prepared teachers and estimated separate value-added models for each variable (also controlling for the covariates in Table 2). Essentially, our first variable indicates whether out-of-state prepared teachers coming from states with more rigorous content standards are more effective than their out-of-state prepared peers coming from states with less rigorous content standards; our second variable tests whether out-of-state prepared teachers coming from states with standards similar to North
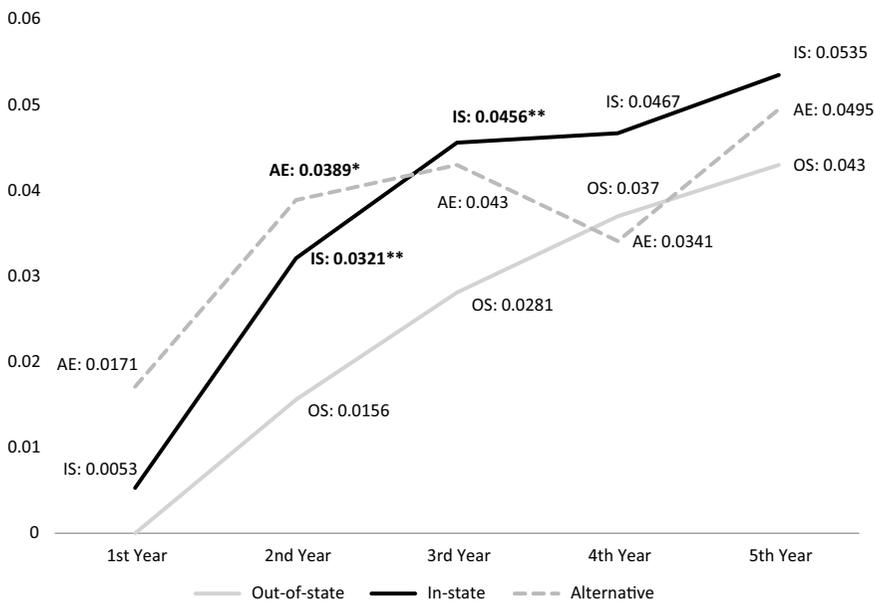
172

FIGURE 3. *Elementary grades reading teacher effectiveness trajectories.*
Note. This figure displays teacher effectiveness relative to first-year out-of-state prepared teachers (coefficients shown in Table 4). For each year of teaching, we used joint significance tests to determine whether there are statistically significant differences between out-of-state prepared and in-state prepared and alternative entry teachers. IS = in-state; AE = alternative entry; OS = out-of-state.
*Statistically significant differences at the *p* < .05 level. **Significance at the .01 level.

Carolina's outperform their out-of-state prepared peers coming from states with dissimilar standards. Examining Table 5, neither of these variables predicted differences in the effectiveness of out-of-state prepared teachers. This is further evidence that a lack of familiarity, specifically a lack of familiarity with the standards and curriculum of the importing state, does not explain the underperformance of out-of-state prepared teachers.

Finally, we performed a third test to determine whether out-of-state teachers prepared in contiguous states—hypothesized to be more familiar with North Carolina's students and educational environment—are more effective than their peers prepared in non-contiguous states. Results from these analyses show no significant differences between the two groups, further reducing confidence that a lack of familiarity explains the lower performance of out-of-state teachers (see Online Appendix Tables 3 and 4, available at http://epa .sagepub.com/supplemental, for these results and a list of states that export the most teachers to North Carolina).

Hypothesis 2: The Quality of Imported Teachers

Background. Prior research indicates that teachers have geographically small labor markets, preferring to work close to their hometown and/ or undergraduate institution (Boyd et al., 2005; Reininger, 2012). In addition, research shows that, on average, teachers with higher levels of human capital—often measured by SAT/ACT scores, licensure exam scores, collegiate selectivity, or grade point average (GPA)—are more effective (Clotfelter et al., 2007, 2010; Dobbie, 2011; Goldhaber, 2007; Greenwald, Hedges, & Laine, 1996; Henry, Bastian, & Smith, 2012; Rockoff, Jacob, Kane, & Staiger, 2011). Taken together, we hypothesize that in states/labor markets where the supply of teachers exceeds demand, if the labor market operates with reasonable efficiency, the teachers with higher levels of human capital (cognitive ability and job-specific knowledge and skills) will be hired locally and the teachers with lower levels of human capital will be forced to broaden their job search to states experiencing teacher shortages.

173

TABLE 4
Teacher Effectiveness Trajectories

	Elementary math		Elementary reading	
	Coefficient	*p* value	Coefficient	*p* value
Second-year teachers	0.0458	.000	0.016	.018
Third-year teachers	0.0765	.000	0.028	.000
Fourth-year teachers	0.0811	.000	0.037	.000
Fifth-year teachers	0.0851	.000	0.043	.000
	Coefficient	Joint significance test, *p* value	Coefficient	Joint significance test, *p* value
In-state	0.0204	.012	0.005	.390
In-state × Second year	0.0145	.000	0.011	.005
In-state × Third year	−0.0003	.014	0.012	.005
In-state × Fourth year	−0.0029	.041	0.004	.134
In-state × Fifth year	0.0055	.003	0.005	.126
Alternative	0.0283	.060	0.017	.159
Alternative × Second year	0.0312	.000	0.006	.019
Alternative × Third year	−0.0034	.083	−0.002	.173
Alternative × Fourth year	−0.0124	.294	−0.020	.803
Alternative × Fifth year	−0.0149	.348	−0.011	.570
Observations used	715,172		1,008,362	

Note. First-year out-of-state prepared teachers are the reference category in these models. The focal variables are interactions between the teacher preparation categories and teacher experience; models include all other covariates listed in Table 2. We used joint significance tests to compare out-of-state prepared teachers with a given level of experience to in-state prepared and alternative entry teachers at the same experience level. We graph these results and mark statistical significance in Figures 2 and 3.

TABLE 5
Familiarity With Similar State Standards

	Rich-covariate adjustment		School fixed effects		School-year fixed effects	
Focal variable	ES math	ES read	ES math	ES read	ES math	ES read
Standards rating difference	0.002	−0.000	0.003	−0.001	0.000	−0.003
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Observations used	257,460	350,841	235,669	311,529	189,386	243,821
Standards rating difference (absolute value)	−0.002	0.001	0.005	−0.000	0.001	−0.002
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Observations used	257,460	350,841	230,886	310,320	180,866	238,916

Note. For this analysis, we limited our sample to out-of-state prepared teachers. We included all student, classroom/teacher, and school covariates from Table 2 in these models.
*Significance at the .05 level. **Significance at the .01 level.

Empirical Tests and Results. To empirically examine our second hypothesis, we performed two tests. First, to determine whether differences in human capital between out-of-state prepared and in-state prepared and alternative entry teachers mediate effectiveness differences between the

174

TABLE 6
Examining the Quality of Imported Instructors

	Rich-covariate adjustment		School fixed effects		School-year fixed effects	
	ES math	ES read	ES math	ES read	ES math	ES read
Do teacher test scores significantly predict student achievement gains?						
Standardized teacher	0.029**	0.005*	0.028**	0.005*	0.026**	0.005
test scores	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Observations used	715,172	1,008,362	714,198	1,006,001	679,625	954,991
Do teacher test scores mediate the teacher preparation effects?						
In-state prepared	0.023**	0.012**	0.019**	0.011**	0.018**	0.009**
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
Alternative entry	0.035**	0.013*	0.034**	0.026**	0.025**	0.026**
	(0.007)	(0.005)	(0.007)	(0.005)	(0.008)	(0.006)
Standardized teacher	0.030**	0.005*	0.029**	0.006*	0.027**	0.006*
test scores	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Observations used	715,172	1,008,362	638,290	883,837	492,354	673,522

Note. The coefficients on standardized teacher test scores in the top panel test whether this measure of human capital significantly predicts student achievement. The coefficients in the bottom panel, when compared with those in Table 3, indicate whether teacher test scores mediate the differences in teacher preparation category effectiveness. We included all student, classroom/teacher (except teacher preparation categories), and school covariates for results in the top panel and all covariates for the results in the bottom panel.
*Significance at the .05 level. **Significance at the .01 level.

categories, we included an indicator of teacher human capital in our adjusted-average value-added models. To select a mediator for this analysis, we required a measure of human capital taken prior to entry into the teaching profession that is significantly associated with student achievement gains. Available mediators included the competitiveness of the undergraduate university attended (Barron's ranking) and standardized teacher exam scores (e.g., SAT/ACT, Praxis I, and Praxis II exams), each of which are significantly associated with teacher value-added in prior work (Clotfelter et al., 2007, 2010; Goldhaber, 2007). However, when implementing the first step of the mediation analyses—determining whether the mediator significantly predicts the outcome— only the teacher exam scores returned significant results (Shrout & Bolger, 2002). We display the coefficients for the teacher exam scores from these first stage models—which controlled for student, classroom/teacher, and school covariates but excluded the teacher preparation categories— in the top panel of Table 6. To examine whether this human capital indicator mediates the overall effectiveness differences shown in Table 3,

we re-ran our rich covariate and school and school-by-year fixed effects models and included the standardized teacher test score variable.[5] Here, if differences in human capital account for the overall effectiveness differences between the groups, we would expect out-of-state prepared teachers to no longer be significantly less effective after the inclusion of the test score variable. Instead, the bottom panel of Table 6 shows no evidence of mediation: Even after controlling for teacher test scores, in-state prepared and alternative entry teachers are significantly more effective and the coefficients remain equivalent in size. Although this test does not examine whether there are human capital differences between out-of-state prepared teachers working in North Carolina and their peers who remained in their states of origin,[6] it does indicate that differences in teachers' exam scores do not explain why out-of-state prepared teachers are less effective.

Second, we examined whether out-of-state prepared teachers coming from states with greater competition for open teaching positions (states with a surplus of teachers) are less effective—as our hypothesis predicts—than their out-of-state

175

TABLE 7
Does Overcrowding Influence Out-of-State Prepared Teacher Performance?

Focal variable	Rich-covariate adjustment		School fixed effects		School-year fixed effects	
	ES math	ES read	ES math	ES read	ES math	ES read
Standardized change in enrollment (2006–2010)	0.003	0.003	0.004	0.003	0.004	0.002
	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)
Standardized % of a state's initially licensed teachers prepared out-of-state (2006–2010)	−0.008	−0.006	−0.011*	−0.009*	−0.010	−0.007
	(0.004)	(0.003)	(0.005)	(0.004)	(0.006)	(0.004)
Observations used	209,050	288,287	193,987	266,042	157,102	215,443

Note. We limited the sample for this analysis to out-of-state prepared teachers only. Change in enrollment is the standardized change in elementary grades (K–5) enrollment from 2005–2006 through 2009–2010 in each state. Percentage of a state's initially licensed teachers prepared out-of-state is the average, over 5 years (2005–2006 through 2009–2010) of the teachers initially licensed in a state who were prepared in another state. We included all student, classroom/teacher (except teacher preparation categories), and school covariates from Table 2 in these models.
*Significance at the .05 level. **Significance at the .01 level.

prepared peers coming from states with less competition for open teaching positions. Because we do not have teacher supply and demand data for each state, we used Common Core and Title II data to construct two additional variables for this hypothesis test: (a) From the Common Core Data, we created a variable that is the percentage change in each states' elementary grades (K–5) student enrollment during our study period (2005–2006 through 2009–2010) and (b) from the Title II Higher Education Act data, we generated a variable that is the average percentage (over our 5-year study period) of each states' initially licensed teachers who were prepared in another state.[7] Adjusting for changes in states' elementary grades student enrollment, we expect teachers entering North Carolina from states that import a larger percentage of their initially licensed workforce to be less effective than teachers coming from other states. This may be the case because (a) teachers prepared in states that import a larger percentage of their workforce face greater competition for open teaching positions and (b) districts and schools may be dissatisfied with the overall quality of in-state prepared teachers (causing them to look to other states for quality teachers). For this analysis, we limited our sample to out-of-state prepared teachers only and merged these two new variables into the data such that the values for a particular out-of-state prepared teacher are linked back to the state in which she was prepared (e.g., teachers prepared in Ohio are assigned the values for Ohio's K–5 student enrollment change and Ohio's percentage of initially licensed teachers prepared in another state).

Results in Table 7 show that, controlling for changes in elementary grades student enrollment, a one standard deviation increase in the percentage of a state's initially licensed teachers who were prepared in another state is associated with significantly less effective mathematics and reading instructors in the school fixed effects specification. Coefficients in both rich-covariate adjustment models and in the mathematics school-year fixed effects model are all significant at the .10 level. In other words, as a state imports a larger percentage of its initially licensed teacher workforce, the early-career teachers who leave that state and come to North Carolina are less effective, on average, than the early-career teachers entering North Carolina from states that import a smaller percentage of their early-career teacher workforce. Assuming this variable captures greater competition for open teaching positions, these test results suggest that those exiting states with a surplus supply of teachers are less effective.

Hypothesis 3: Teacher Turnover and Differential Attrition

Background. If teachers have strong preferences to teach close to home but are forced to seek

176

teaching positions in other states, we hypothesize that as these out-of-state prepared teachers acquire more job-specific knowledge and skills through experience teaching, they will be become more attractive for open teaching positions back in their states of origin (Boyd et al., 2005; Reininger, 2012). If this is true, out-of-state prepared teachers can be expected to leave North Carolina public schools at significantly greater rates than in-state prepared teachers. This turnover may adversely influence student achievement, reduce school stability, increase teacher recruitment and training costs, and contribute to large numbers of novice teachers in the workforce (Alliance for Excellent Education, 2004; Henry, Bastian, & Fortner, 2011; Henry, Fortner, & Bastian, 2012; Ronfeldt, Loeb, & Wyckoff, 2013). In addition to the potential adverse effects of high turnover rates, differential attrition may contribute to out-of-state prepared teachers' ineffectiveness (Boyd et al., 2006). It is important to note, however, that the effect of differential attrition on preparation category estimates is influenced by the number and effectiveness of exiting out-of-state prepared, in-state prepared, and alternative entry teachers.

Empirical Tests and Results. To evaluate our final hypothesis, we perform two tests. First, to determine whether out-of-state prepared teachers leave North Carolina public schools at significantly greater rates than in-state prepared and alternative entry teachers, we modeled teacher turnover with a logistic regression equation. Here, exiting the state's public schools is the dependent variable (not returning the following school year),[8] out-of-state prepared teachers are the reference category, indicators for in-state prepared and alternative entry teachers are the focal variables, and we include a set of classroom and school covariates (the same as those listed in Table 2) to control for differences in employment context that may influence persistence in North Carolina public schools. The equation for this specification is as follows[9]:

$$\Pr(Exit_{it}=1)=\frac{\exp\left(teacherprep_i\alpha+Class_{it}\beta+School_{it}\tau\right)}{1+\exp\left(teacherprep_i\alpha+Class_{it}\beta+School_{it}\tau\right)},\quad (2)$$

where $Exit_{it}$ is a binary indicator for whether or not a teacher returns to North Carolina public schools in the following school year; $teacherprep_i$ is a set of teacher preparation indicators for in-state and alternative entry

TABLE 8
Logistic Regression Results for Teacher Turnover

Teacher category	Odds ratio and *Z* score
In-state prepared	0.464**
	(−14.48)
Alternative entry	1.291**
	(3.62)
Observations used	22,910

Note. Odds ratios are in relation to out-of-state prepared teachers, with teachers exiting North Carolina public schools as the dependent variable. Observations identify unique teacher-year combinations.
*Statistical significance at the .05 level. **Significance at the .01 level.

instructors in reference to out-of-state prepared teachers; and $class_{it}$ and $school_{it}$ are a set of classroom/teacher and school contextual factors that may influence teacher persistence.

Results in Table 8 show that in-state prepared teachers have approximately one half the odds of exiting North Carolina public schools as their out-of-state prepared peers, whereas alternative entry teachers have significantly greater odds of exiting teaching in the state's public schools. During our study period, 25% of the early-career out-of-state prepared teachers in our sample exited North Carolina public schools; only 13% of in-state prepared teachers left the state's public schools. These high rates of turnover are consistent with the hypothesis that as out-of-state prepared teachers acquire human capital through teaching experience in North Carolina they become more competitive for open teaching positions in their states of origin.

Second, to determine whether differential attrition explains the performance of out-of-state prepared teachers, we examined the effectiveness of teachers who depart versus those who stay. To do so, we followed previous research and employed a two-stage approach where in Stage 1, we estimated individual teacher-year effectiveness, using the same model and covariates used to estimate individual teacher-year effectiveness for Research Question 1. Next, in Stage 2, we made the individual teacher-year effectiveness estimate our outcome measure and specified a model with the following focal covariates: indicators for in-state prepared and alternative entry teachers (in reference to

TABLE 9

Differential Attrition for Out-of-State Prepared Teachers

Focal variables	ES math	ES reading
In-state prepared	0.018**	0.007*
	(0.004)	(0.003)
Alternative entry	0.026**	0.007
	(0.008)	(0.007)
Will leave	−0.035**	−0.024**
	(0.007)	(0.005)
In-state × Will leave	−0.007	0.008
	(0.010)	(0.008)
Alternative entry × Will leave	0.016	0.013
	(0.016)	(0.013)
Cases	22,078	22,438

Note. To examine differential attrition, we used a two-stage approach. In Stage 1, we estimated individual teacher-year effectiveness controlling for student, classroom, and school characteristics (excluding teacher characteristics). Then, for Stage 2, we made individual teacher-year effectiveness the outcome measure and included the covariates shown above (in reference to out-of-state prepared teachers) in a regression. *Significance at the .05 level. **Significance at the .01 level.

out-of-state prepared teachers), will leave (a time-invariant variable equal to one for those who will exit North Carolina public schools), and interactions between will leave and in-state prepared and alternative entry teachers.[10] Here, the main effect of will leave captures whether out-of-state prepared teachers who were more or less effective during their tenure in North Carolina public schools leave the state, while the interaction variables indicate whether the effectiveness of exiting in-state prepared or alternative entry teachers significantly differs from the effectiveness of exiting out-of-state prepared teachers.

Results in Table 9 show that exiting out-of-state prepared teachers are significantly less effective than their out-of-state prepared peers who stay—by 0.035 and 0.024 standard deviations in mathematics and reading, respectively. This result is consistent with recent research evidence which indicates that, on average, less effective teachers are more likely to exit teaching (Feng & Sass, 2011; Goldhaber et al., 2011; Henry, Bastian, & Fortner, 2011). Furthermore, the interaction coefficients in Table 9 are statistically insignificant, meaning that (a) exiting in-state prepared and alternative entry teachers

perform similarly to exiting out-of-state prepared teachers and (b) the overall effectiveness differences shown in Table 3 would not necessarily narrow due to the exit of less effective out-of-state prepared teachers.

Discussion

In response to a need for both more and better teachers, states have created alternatives to solely licensing instructors prepared at in-state traditional education programs. While alternative entry routes have garnered the most policy and research attention during this time, many states have also reduced barriers to employment and broadened their potential labor pool of traditionally prepared instructors by granting certification for out-of-state prepared teachers. Collectively, these reciprocal certification policies have helped facilitate the interstate movement of teachers and have aided high-growth states meet the demand for more teachers. Until recently, however, the effects of this policy choice were largely unexplored. Therefore, in this study, we separated traditionally prepared teachers into two groups—in-state and out-of-state prepared—and assessed both the comparative effectiveness of out-of-state prepared teachers and potential explanations for differences in out-of-state prepared teacher performance.

Results indicated effectiveness differences between the teacher preparation categories: In-state prepared and alternative entry teachers significantly outperformed their out-of-state prepared peers across all model specifications in elementary grades mathematics and reading. Beyond statistical significance, these results take on practical significance because out-of-state prepared teachers are the second largest source of early-career instructors in North Carolina elementary schools (more than one third of our sample). Given these results and the possibility that these findings could be different elsewhere, education policy researchers may wish to perform similar analyses to assess the effects of out-of-state prepared teachers in other states. Furthermore, future analyses may address a limitation of the present study by creating more refined teacher categories that account for teachers' education experiences prior to their preparation. This is important because the current study

178

may miss heterogeneity in the effectiveness and persistence between those out-of-state prepared teachers with, and those without, prior schooling experiences in North Carolina.

Moving forward with effective policy solutions requires an understanding of the distribution of teacher effectiveness and why out-of-state prepared teachers struggle. Here, our teacher effectiveness graphs (Figure 1) illustrate a significant overlap in the distribution of teacher effectiveness—there are many high-performing out-of-state prepared teachers, and conversely, many low-performing in-state prepared teachers. Therefore, blunt policy instruments, such as eliminating reciprocal certification, that may block many highly effective out-of-state prepared teachers and cause further teacher shortage concerns are not effective responses.

Findings from our hypotheses tests show that initial differences in effectiveness between in-state prepared and out-of-state prepared teachers remain constant or widen over the early-career period. Furthermore, there is significant across-state heterogeneity in teacher effectiveness, as states with greater competition for open teaching positions export less effective teachers to North Carolina. Taken together, these results indicate that differences in human capital help explain the underperformance of out-of-state prepared teachers. More importantly, this variation in teacher effectiveness across states suggests the potential for state policymakers and local hiring officials to use research on the average effectiveness of teachers from other states to improve teacher recruitment and hiring practices. For example, among the 10 states from which North Carolina imports the most teachers (see Online Appendix Table 4, available at http://epa.sagepub.com/supplemental), teachers coming from New York, Michigan, South Carolina, and West Virginia perform comparably to in-state prepared teachers in both mathematics and reading, whereas teachers coming from Pennsylvania, Ohio, and Virginia are significantly less effective in both subjects. Because of the local nature of teacher labor markets, researchers could go further and identify from which states particular regions or school districts are drawing high- and low-performing teachers.

In addition, results from our hypotheses tests indicate the importance of states and school districts considering both teacher effectiveness *and* teacher persistence when making policy (licensure) or recruitment/hiring decisions. As expected, early-career out-of-state prepared teachers have significantly greater odds (approximately twice as high) of exiting North Carolina public schools than their in-state prepared peers. From a policy perspective, this result is troubling—because out-of-state prepared teachers comprise more than one third of early-career elementary grades teachers and turnover has adverse effects on student achievement, school budgets, school stability, and precipitates the entry of large numbers of novice teachers—and suggests that although effectiveness differences between in-state and out-of-state prepared teachers are small (see Figure 1), school districts may have multiple reasons to prefer in-state prepared teachers (Alliance for Excellent Education, 2004; Ronfeldt et al., 2013). Furthermore, results showed that exiting out-of-state prepared teachers were significantly less effective during their tenure in North Carolina public schools than their out-of-state prepared peers who stay. Although this indicates that differential attrition is not a factor in the underperformance of out-of-state prepared teachers, it again emphasizes the need for communication between researchers and states/school districts. For example, teachers from South Carolina, New York, and West Virginia are not only as effective as in-state prepared teachers but also persist in North Carolina public schools at high rates (approximately 80% stay throughout our study period).

Overall, our results highlight the importance of teacher labor markets to student achievement and the composition of the teacher workforce. Further evidence on the difficulty of attracting highly effective teachers from out-of-state comes from North Carolina's two largest school districts—Wake County and Charlotte-Mecklenburg. Based on labor market theory, we expect these two urban, economically vibrant areas with better amenities and larger teacher salary supplements to recruit/attract higher quality teachers (in-state, out-of-state, and alternatively prepared) than North Carolina's rural regions (Glaeser & Mare, 2001). Consistent with this expectation, we find that out-of-state prepared teachers in Wake County and Charlotte-Mecklenburg are significantly more effective

than their out-of-state prepared peers in the state's other 113 school districts. However, out-of-state prepared teachers in Wake County and Charlotte-Mecklenburg are significantly less effective than the in-state prepared (mathematics only) and alternative entry teachers working in those districts (see Online Appendix Table 6, available at http://epa.sagepub.com/supplemental). Therefore, even in well-resourced areas, the differential between in-state and out-of-state prepared teachers is maintained.

Beyond the need for researchers to provide evidence on the effectiveness and attrition of teachers from out-of-state to facilitate evidence-based policy and hiring, the results of these analyses suggest two additional solutions. First, given the performance and persistence of out-of-state prepared teachers, states may choose to build and retain their own labor supply. Here, approaches might include providing more competitive salaries for early-career teachers,[11] increasing the recruitment and production goals for in-state preparation programs, and improving the yield—the percentage of graduates hired in-state—for in-state preparation programs. To retain those already in the profession, states can develop and rigorously evaluate programs for induction, instructional coaching, and improving the quality of school working conditions and administrators. Second, because there were many low-performing teachers from each category (see Figure 1), states/districts should develop better procedures for identifying applicants who are most likely to be effective teachers. Here, potential practices that could be developed and evaluated include assessing job candidates' performance while teaching a sample lesson or using measures of candidates' non-cognitive characteristics (e.g., grit, organizational skills, leadership) to better recruit and hire teachers (Bastian, 2013; Duckworth, Quinn, & Seligman, 2009; Rockoff et al., 2011).

Given the current labor market context, where individuals are more mobile and change positions more frequently, understanding the potential effects of greater teacher workforce mobility is critical. Our work suggests that out-of-state prepared teachers help address teacher shortages in the importing state, but, on average, are less effective and more likely to leave the importing

state's public school classrooms. Therefore, we contend that when importing teachers to solve supply issues, states and districts must use research evidence to make hiring decisions in light of labor market factors.

Notes

1. Approximately 2% of the sample was unclassifiable because (a) they did not have a college graduation year in the administrative data; (b) their highest degree earned prior to beginning teaching was less than a bachelor's degree; or (c) administrative data recorded the person teaching more than 1 year prior to their earliest graduation year. We retained these unclassifiable teachers in analyses but do not report their effects. Results for unclassifiable teachers are available on request.

2. We get the prior score for third-grade students from the North Carolina Grade 3 pre-test. The state stopped administering this pre-test after the 2008–2009 school year, meaning third-grade value-added outcomes are not estimated in 2009–2010.

3. These effectiveness findings for alternative entry teachers are robust to model specifications that separate two high-performing groups of teachers—Teach For America and Visiting International Faculty—from the much larger general category of alternative entry teachers.

4. Through the recent adoption of the Common Core State Standards, states' standards and assessment systems will soon become more similar. During

the study period for this analysis (2005–2006 through 2009–2010), however, there were significant differences between states.

5. Approximately 18% of the teachers in our study sample were originally missing a value for the composite teacher test score variable. To run our mediation analyses on the same sample of teachers and students, we created a data set with all available information about our sample teachers and the classrooms and schools in which they work and used the SAS proc mi (multiple imputation) command to impute the missing teacher test score values. Findings from a non-imputed mediation analysis are comparable with the presented results and are available on request.

6. We cannot perform this test because we do not have access to indicators of human capital—exam scores or undergraduate selectivity—for out-of-state prepared teachers who remain in their states of origin.

7. During our study period, New York State did not report the percentage of its initially licensed teachers who were prepared in another state. Therefore, we treat this data as missing and teachers from New York are excluded from this hypothesis test.

8. Certified salary files provided by the North Carolina Department of Public Instruction allow us to know whether or not a teacher remains in the state's public schools. If a teacher exits the state's public schools, available data do not allow us to determine whether she takes a teaching position in a private school or out-of-state or exits the teaching profession.

9. In addition to this logistic regression, we also performed logistic regressions with school and school-by-year fixed effects. Results from these models were very similar to our findings presented in Table 8.

10. In Online Appendix Table 5 (available at http://epa.sagepub.com/supplemental), we present results from another two-stage differential attrition analysis. Stage 1 remains the same, but for Stage 2, we limited our sample to out-of-state prepared teachers, made will leave the outcome measure, and regressed will leave on quintiles of teacher effectiveness (with quintile 1, the lowest performing, as the reference group). For out-of-state prepared teachers results indicate that (a) the less effective have significantly greater odds of exiting North Carolina public schools and (b) there is no evidence of a bi-modal—both the most and least effective—attrition pattern.

11. In the 2011–2012 school year, the average beginning teacher salary in North Carolina was US$30,779—a ranking of 46th nationally (National Education Association, 2012). Therefore, issues of teacher compensation may be particularly important to recruit and retain higher quality early-career teachers.

References

Alliance for Excellent Education. (2004). *Tapping the potential: Retaining and developing high-quality teachers*. Retrieved from http://all4ed.org/wp-content/uploads/2007/07/TappingThePotential.pdf

Bacolod, M. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics*, *89*, 737–751.

Bastian, K. C. (2013, March 14). *Do teachers' non-cognitive skills and traits predict effectiveness and evaluation ratings?* Paper presented at the Association for Education Finance and Policy 38th Annual Conference. Retrieved from https://aefp-web.org/sites/default/files/webform/AEFP_non-cognitive.pdf

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *The Future of Children*, *17*, 45–68.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, *1*, 176–216.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, *31*, 416–440.

Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools. *Journal of Policy Analysis and Management*, *30*, 88–110.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, *24*, 113–132.

Carmichael, S., Martino, G., Porter-Magee, K., & Wilson, W. S. (2010). *The state of state standards—and the Common Core—in 2010*. Thomas Fordham Institute. Retrieved from http://edexcellence.net/sites/default/files/publication/pdfs/SOSSandCC2010_FullReportFINAL_8.pdf

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, *24*, 305–331.

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*, 673–682.

181

Clotfelter, C., Ladd, H., & Vigdor, J. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*, 655–6861.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, *23*, 57–77.

Dobbie, W. (2011). *Teacher characteristics and student achievement: Evidence from Teach For America*. Retrieved from http://blogs.edweek .org/edweek/teacherbeat/teachercharacteristics-july2011.pdf

Duckworth, A., Quinn, P., & Seligman, M. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, *4*, 540–547.

Feistritzer, C. E. (2011). *Profiles of teachers in the U.S. 2011*. Washington, DC: National Center for Education Information. Retrieved from http://www.edweek.org/media/pot2011final-blog.pdf

Feng, L., & Sass, T. (2011). *Teacher quality and teacher mobility* (Calder Institute Working Paper No. 57). Retrieved from http://www.caldercenter .org/upload/CALDERWorkPaper_57.pdf

Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, *19*, 316–342.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, *42*, 765–794.

Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*, 129–145.

Goldhaber, D., Gross, B., & Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management*, *30*, 57–87.

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*(2), 29–44.

Greenberg, J., Pomerance, L., & Walsh, K. (2011). *Student teaching in the United States*. National Council on Teacher Quality. Retrieved from http://www.nctq.org/dmsView/Student_Teaching_United_States_NCTQ_Report

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*, 361–396.

Hanushek, E., Kain, J., & Rivkin, S. (2004). Why public schools lose teachers. *Journal of Human Resources*, *39*, 326–354.

Henry, G. T., Bastian, K. C., & Fortner, C. K. (2011). Stayers and leavers: Early-career teacher effectiveness and attrition. *Educational Researcher*, *40*, 271–280.

Henry, G. T., Bastian, K. C., Fortner, C. K., Kershaw, D. C., Purtell, K. M., Thompson, C. L., & Zulli, R. A. (in press). Teacher preparation policies and their effects on student achievement. *Education Finance and Policy*.

Henry, G. T., Bastian, K. C., & Smith, A. A. (2012). Scholarships to recruit the "best and brightest" into teaching: Who is recruited, where do they teach, how effective are they, and how long do they stay? *Educational Researcher*, *41*, 83–92.

Henry, G. T., Fortner, C. K., & Bastian, K. C. (2012). The effects of experience and attrition for novice high-school science and mathematics teachers. *Science*, *335*, 1118–1121.

Henry, G. T., Thompson, C. L., Fortner, C. K., Bastian, K. C., & Marcus, J. V. (2011). *Technical report: UNC teacher preparation program effectiveness*. Education Policy Initiative at Carolina. Retrieved from http://publicpolicy.unc.edu/files/2014/02/ProgramEffect_TechnicalRpt_July2011.pdf

Ingersoll, R., & Merrill, L. (2010). Who's teaching our children? *Educational Leadership*, *67*(8), 14–21.

Ingersoll, R., & Smith, T. (2003). The wrong solution to the teacher shortage. *Educational Leadership*, *60*(8), 30–33.

Kane, T., Rockoff, J., & Staiger, D. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*, 615–631.

National Education Association. (2012). *NEA collective bargaining/member advocacy database*. Retrieved from http://www.nea.org/home/2011-2012-average-starting-teacher-salary.html

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. Washington, DC: National Academies Press.

North Carolina Department of Public Instruction. (2011). *North Carolina Mathematics Tests. 3rd ed.: Technical report*. Raleigh, NC. Retrieved from http://www.dpi.state.nc.us/docs/accountability/reports/mathtechmanualdrafted2.pdf

Porter, A. C., Polikoff, M. S., & Smithson, J. (2009). Is there a de facto national intended curriculum? Evidence from state content standards. *Educational Evaluation and Policy Analysis*, *31*, 238–268.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teacher's location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis*, *34*, 127–145.

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, *6*, 43–74.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*, 4–36.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422–445.

U.S. Department of Education. (2010). *The secretary's seventh annual report on teacher quality*. Retrieved from http://www2.ed.gov/about/reports/annual/teachprep/t2r7.pdf

Authors

KEVIN C. BASTIAN is a research associate and the director of the Teacher Quality Research Initiative in the Department of Public Policy and the Education Policy Initiative at Carolina at the University of North Carolina at Chapel Hill. His research interests include education policy and evaluation and teacher and school leader preparation, labor markets, and effectiveness.

GARY T. HENRY holds the Patricia and Rodes Hart Chair and serves as a professor of public policy and education in the Department of Leadership, Policy and Organization, Peabody College, Vanderbilt University. His research interests include teacher quality, education policy, educational reform, quantitative research methods, and evaluation.

183