Knocking on the door to the teaching profession? Modeling the entry of prospective teachers into the workforce

Dan Goldhaber [a], John Krieg [b], Roddy Theobald [a,*]

[a] Center for Education Data & Research, University of Washington Bothell, 3876 Bridge Way N, Suite 201, Seattle, WA 98103, United States
[b] Western Washington University, Department of Economics, Parks Hall 315, Bellingham, WA 98225, United States

ARTICLE INFO

ABSTRACT

We use a unique longitudinal sample of student teachers ("interns") from six Washington state teacher training institutions to investigate patterns of entry into the teaching workforce. We estimate split population models that simultaneously estimate the impact of individual characteristics and student teaching experiences on the timing and probability of initial hiring as a public school teacher. Not surprisingly, we find that interns endorsed to teach in "difficult-to-staff" areas are more likely to find employment as public school teachers than interns endorsed in other areas. Younger interns, white interns, and interns who completed their student teaching in suburban schools are also more likely to find a teaching job, all else equal. Prospective teachers who do their internships at schools that have more teacher turnover are more likely to find employment, often at those schools. On the other hand, few of the characteristics of an intern's cooperating teacher are predictive of workforce entry. Finally, interns with higher credential exam scores are more likely to be hired by the school where they did their student teaching.

© 2014 Elsevier Ltd. All rights reserved.

In an era of heightened concern about the economic returns on investments in teacher preparation (or, for that matter, in other higher education programs), data on hiring and placement of teachers is a legitimate component of a broader evaluation of teacher preparation program (TPP) quality. What are the job prospects of TPP graduates? (Anderson & Stillman, 2013)

1. Introduction

The past 20 years have seen a proliferation of empirical research into the composition and distribution of the teacher workforce. Extensive quantitative work investigates where teachers choose to teach, and the factors that determine whether and when teachers choose to leave the public teaching workforce.[1] But there is far less research on the first step of a teacher's career path: who enters the teaching workforce in the first place?

The scarcity of empirical research on entry into the teacher workforce is surprising. Teacher training has come under increased scrutiny (e.g. Greenberg, McKee, & Walsh, 2013), and a growing literature investigates the impact of

---

* Corresponding author. Tel.: +1 206 547 5585; fax: +1 206 547 1641.
E-mail addresses: dgoldhab@uw.edu (D. Goldhaber),
john.krieg@wwu.edu (J. Krieg), roddy@uw.edu (R. Theobald).

---

[1] For instance, for research on where teachers choose to teach see Boyd et al. (2005, 2011, 2013), Maier and Youngs (2009), Reininger (2012); and for research on attrition from the public school labor force see Brewer (1996), Boyd et al. (2008), DeAngelis and Presley (2011), Goldhaber, Gross, and Player (2011), Ingersoll et al. (2012), Krieg (2006), Ronfeldt (2012), Scafidi, Sjoquist, and Stinebrickner (2007), and Stinebrickner (2001).

pre-service training—either the training program itself (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Goldhaber, Liddle, & Theobald, 2013; Koedel, Parsons, Podgursky, & Ehlert, 2014; Mihaly, McCaffrey, Sass, & Lockwood, 2013) or student teaching experiences (Boyd et al., 2006; Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2008)—on teacher mobility and effectiveness. These studies, however, focus on individuals who decided to enter the teaching workforce and received a teaching job. Many studies do address the factors that influence the decision to get a teaching degree or the decision to enter the teaching workforce,[2] but lack detailed information about teacher training experiences, student teaching in particular. As such, the existing literature ignores the potential differential effects of pre-service training experiences and intern characteristics on the probability of workforce entry and outcomes after workforce entry.

In this paper, we provide the first quantitatively descriptive look at the process that moves teachers from training programs and student teaching placements into the teaching workforce. Specifically, we focus on the teacher training experiences of "interns" (i.e., students in traditional teacher training programs who complete student teaching and other requirements to receive a teaching credential) from a sample of six training institutions in Washington State. These interns are linked with longitudinal data to allow us to estimate the probability that individuals who obtain a teaching credential end up employed in a public school teaching job, employed in a private school teaching job, employed in a public school non-teaching job, or not employed in any public or private school job within the state.

After investigating the movement of interns into and between these four outcomes, we consider hiring into a public school teaching position or not being hired at all as a binary outcome (dropping the small number of interns in either private school teaching jobs or public school non-teaching jobs), and estimate split population models that simultaneously model the impact of covariates on the timing and probability that an intern finds a public school teaching job.

Controlling for differences in placement rates by training institution and over time, we find that interns endorsed to teach in "difficult-to-staff areas" like math and science (STEM), special education, and English Language Learning (ELL) are far more likely to be employed in public schools (and are employed more quickly) than interns endorsed in other areas. This is also true for younger and white interns. We find little evidence that characteristics of an intern's cooperating teacher are predictive of entry into the public school workforce.

There is, however, evidence that the type of school in which internships occur matters. Interns who complete their student teaching in a suburban school are more likely to enter public school teaching, as are those who do their

student teaching in a school with high teacher turnover. This finding on teacher turnover is related to the fact that a surprising number of interns, just over 15 percent of all interns hired into an in-state public teaching position, are *hired into the same school where they did their student teaching.* This is an important and novel finding as it suggests that student teaching serves not merely as a means of training teachers but also as a way for schools with open positions to get an early look at prospective teachers, screening them for fit and ability. In fact, when we investigate this specific type of hiring, modeling the probability that an intern is hired into his or her internship school (as opposed to a different school), we find results that differ from those of the hiring model more generally. For instance, interns with higher credential exam scores are more likely to be hired into their internship school, but are not generally more likely to be employed in public schools.

Our analysis unifies and builds on three strands of the teacher labor market literature: impacts of teacher training and student teacher experiences; recruitment and retention of teachers in difficult-to-staff subject areas; and evidence on teacher workforce entry. We discuss these strands of the literature and provide some context to Washington State in Section 2, describe our data in Section 3, give an overview of our analytic approach in Section 4, and then present our results in Section 5. We conclude with some policy implications in Section 6.

2. Background and context

Pre-service training is seen as a process that is fundamental to influencing the over three million teachers currently employed in the K-12 workforce. Chief among these pre-service experiences is student teaching; as Anderson and Stillman (2013) note, "policymakers and practitioners alike increasingly tout clinical experiences as a key component—even 'the most important' component of—pre-service teacher preparation." While a large literature exists on the impacts of student teaching (see Anderson and Stillman (2013) for a comprehensive review), the vast majority of these studies are case studies with very small sample sizes.[3]

A small quantitative literature uses substantially larger samples to link various features of teacher training to data on teacher career paths and effectiveness.[4] Boyd et al. (2006) find evidence that programs that include a capstone project—where teachers relate curriculum learning to actual practices—as part of the student teaching experience tend to produce more effective first-year teachers. Boyd et al. (2008) find that, in terms of students' math achievement in particular, teachers who identify similarities between their student teaching experience and their

first-year classroom experiences have greater student achievement gains. More recently, Ronfeldt (2012) suggests teacher pre-service placement may be linked both to the length of time a teacher stays in the school district and to teacher value-added gains in student achievement, while Ingersoll, Merrill, and May (2012), Papay, West, Fullerton, and Kane (2012), and Ronfeldt, Schwartz, and Jacob (2014) each find positive effects of more extensive teacher training on teacher retention.

An important shortcoming of the literature described above is that it focuses exclusively on a sample of individuals who are already in the teaching workforce, and thus ignores the possibility that pre-service might affect the likelihood that individuals are hired as teachers. This could mean that the existing literatures provide a misleading picture of the efficacy of training practices. For instance, suppose that a particular pre-service training intervention is found to positively impact the effectiveness and retention of those individuals who enter the workforce, but negatively affects the likelihood that prospective teachers opt to enter the profession. It is conceivable that the benefits of the intervention for in-service teachers are offset by the increased cost associated with having to train more people for a comparable yield. Clearly the yield of teacher trainees must be considered as part of an analysis of the cost-effectiveness of any pre-service intervention.

However, there is little empirical evidence on the factors that influence teacher workforce entry. A few studies investigate the differences between college graduates who do and do not enter teaching, finding that college students who opt to go into teaching tend to be less academically proficient as measured, for instance, by college entrance exams (Goldhaber & Liu, 2003; Hanushek & Pace, 1995; Podgursky, Monroe, & Watson, 2004).[5] There is also evidence (e.g. Bacolod, 2007, Goldhaber & Liu, 2003; Ingersoll & Perda, 2010) that graduates with degrees in science, technology, engineering, and math (STEM) areas are less likely to become teachers.

Each of these studies compares individuals who decide to become teachers with college graduates or attendees who decide not to become teachers, but this is not the relevant comparison group for all policy questions. Specifically, if we are interested in the impacts of teacher training experiences, training programs cannot have an impact on students who do not enroll in their programs. Likewise, if we are interested in school hiring practices, schools cannot hire teachers who do not have a teaching degree. Thus we argue that the relevant comparison group, at least in these cases, is individuals who *did* get a teaching degree but *did not* become a teacher.

To our knowledge, only four papers have focused on the transition of *prospective teachers* from training programs into the teaching workforce. Ballou (1996) focuses on the school side of the teacher hiring process, and finds little evidence that strong academic credentials help a prospective teacher's job prospects. Engel, Jacob, and Curran
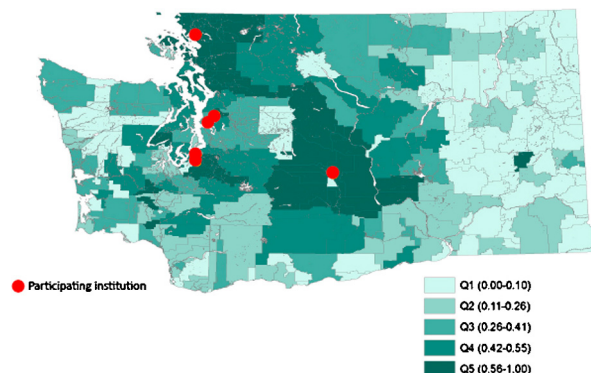

Fig. 1. Proportion of new teachers from participating institutions.

(2014), on the other hand, focus on the preferences of prospective teachers (as measured by the schools where they choose to apply), and find that schools serving more advantaged students receive more applicants per vacancy. Boyd, Lankford, Loeb, and Wyckoff (2013) use a two-sided matching model to try to disentangle the preferences of teachers and schools. Their findings run contrary to Ballou (1996) in that they do find evidence that schools demonstrate preferences for prospective teachers with stronger academic credentials, and reinforce the conclusion from Engel et al. (2014) that prospective teachers prefer schools with more advantaged students. Finally, White, DeAngelis, and Lichtenberger (2013) follow prospective teachers from Illinois colleges into the Illinois public teaching workforce and find that prospective teachers who are non-white are less likely to be employed as a public teacher teachers than those who are white even though whites and non-whites in Illinois are about equally likely to earn their teaching certificate.

3. Data and descriptive statistics

3.1. Data sources

We link information from two sources: teacher training institutions (TTIs), and Washington's Office of the Superintendent of Public Instruction (OSPI). We received data from six Washington State TTIs—Central Washington University, Pacific Lutheran University, University of Washington-Bothell, University of Washington-Seattle, University of Washington-Tacoma, and Western Washington University—about college students who completed student-teaching internships in the state's public schools. Five of the six universities are located in the western third of the state and none are in the eastern third so it is not surprising that the institutions in our sample of disproportionately serve school systems on the western side of the state (see Fig. 1, showing the percentage of teachers in each school district in the state that were trained in these TTIs).

TTIs provided information on each college student who completed a student–teacher internship (referred to as "interns") during a specific range of years, though the

---

[5] See Goldhaber and Walch (2014) for preliminary evidence that the academic proficiency of new teachers (relative to non-teachers) may be improving.

range of years for which data were available varies by TTI.[6] TTIs also provided the academic year of the internship, the building and district in which the internship occurred, and the name of the teacher supervising the internship (the "cooperating" or "mentor" teacher).[7] Some universities also provided additional demographic and extended academic background data about their interns.[8] The TTIs in our sample graduate roughly one third of the teachers who enter the Washington state teaching workforce each year, and include three of the four largest teacher training institutions in the state (as measured by the average number of workforce entrants from each program).[9]

We merge the TTI data with administrative data provided by OSPI containing annual information about employment, years of experience, race, and educational background for every K-12 public school employee in the state between 1994 and 2011 and every private school teacher from 2004 to 2011,[10] as well as endorsements (the training specialty recognized by the state) for all individuals who are credentialed before 2013.[11] We merge the above sources of data to *both* interns and their cooperating teachers, creating a dataset that tracks whether interns were hired into Washington state public schools (either as teachers or in a non-teaching role) or private schools and links these interns to their cooperating teachers. Unfortunately we do not know whether interns who are not eventually employed in Washington's public or private schools might have found employment outside of K-12 education in Washington or in any non-K-12 employment inside of Washington.

In addition to individual-level information on interns and their supervisors, we make use of annual OSPI data that describe the school at which the internship takes place. These data include total enrollment, the percent of students who pass the state math and reading exams,[12] the percent of students by federal ethnicity categories, the percent of students enrolled in the free/reduced lunch program, the location of the school (urban, suburban, town, or rural), and whether the school is in a district that shares a border with Oregon, Idaho, or Canada. We compute the number of prior interns that we observe to have completed their internships at each school, which provides a rough measure of each school's experience with student teachers.

Following Ronfeldt (2012), we use OSPI employment data to calculate the "stay ratio" of each internship school, which is a measure of teacher turnover.[13] We modify Ronfeldt's definition and define a school's stay-ratio in a given school year as the percent of the school's non-retirement-age teachers who return to the school in the following year.[14] Therefore, schools with *less* teacher turnover have a *higher* stay ratio. We also use this longitudinal dataset to create an indicator for whether each intern's student teaching school hires a new (to the school) teacher the following school year. This is important because a large percentage of interns (about 15%) are hired into the same school where they did their student teaching.

An important variable that we observe for most, but not all, of the interns in our final sample is ethnicity. We compile intern ethnicity from three sources: the OSPI administrative data (which contains all hired interns, with a small amount of missing ethnicity data); the endorsement file (which contains all interns, but with a considerable amount of missing ethnicity data); and the dataset of interns from Western Washington University (with no missing ethnicity data). From these three sources, we are able to create ethnicity indicators (American Indian, Asian, black, Hispanic, or white) for 94% of the interns in our sample.[15] For parsimony, we create a binary variable indicating whether each intern is non-white. In our

---

[6] The longest span provided by a university was every intern between the years 1998 and 2011 and the shortest span was 2006–2011.

[7] 616 interns (all from Western Washington, UW-Bothell, or UW-Tacoma) completed more than one internship. Representatives from these universities report that an intern's first internship is often for observational purposes, while the second is when he/she does student teaching. So for these interns, we include the intern's second internship experience in our final dataset. A very small number of interns from Western Washington University completed two student teaching internships. For these interns, we randomly select one internship experience to include in our analytic dataset.

[8] These include high school information (school attended, class standing, and GPA), standardized test scores (SAT and ACT), collegiate GPA, and demographic information (first generation college student and detailed race/ethnic codes).

[9] There are a total of 20 TTIs in Washington (see Goldhaber et al. (2013) for a full list.) Approximately 15 percent of the state's public school teachers were trained outside the state (see Table 1). See http://program.pesb.wa.gov/reports/reporting_progress/clinicallocation for detailed maps on where Washington teachers tend to do their student teaching.

[10] When representing years, this paper uses the convention of listing the first year of the academic year. Thus, 1994 represents the 1994–1995 academic year.

[11] We combine specific endorsement information into five categories: elementary education, special education, STEM, ELL (English Language Learners), and other. The PESB data also contains the birth year of each intern, which allows us to calculate the age of each intern during his or her internship year. We do not observe degree level for non-hired interns in our sample. For hired interns, interns entering the workforce with a masters degree are on average older (31.5 years) than interns entering the workforce with a bachelor's degree (29.2 years).

[12] Although Washington state now tests all students in math and reading in grades 3–8 and 10 each year, for many years in our sample the state only tested students in grades 4, 7, and 10. To calculate the percent of students passing the state exam in math and reading for each year, we first select the grade in each school (4th, 7th, or 10th) in which the most students took the state exam, and then calculate the percent of students who passed the test in that grade. We standardize these passing rates by grade and year to control for differences in the difficulty of the exams in different grades and years.

[13] Ronfeldt (2012) shows that a school's stay ratio is correlated with other survey-based measures of school functionality, such as administrative quality, staff support, student behavior, and teacher safety.

[14] We follow Ronfeldt (2012) by transforming the stay ratios with an exponential transformation and standardizing within school level (elementary or secondary). Ronfeldt uses an average of each school's stay ratio over the 5-year span of his data, and we experiment with several moving averages, including a 3 year moving average (the current year and two prior years) and two 5-year moving averages (the current year and four prior years, and the current year, two prior years, and two subsequent years). Our results use the 5-year moving average calculated over the current year and four previous years, but the results are robust to the choice of average.

[15] In 2011, 3.5% of teachers in Washington were Hispanic, 2.5% were Asian, 1.3% were black, and 1.0% were American Indian. Among interns in our sample for whom we observe ethnicity, 2.9% are Hispanic, 4.4% are Asian, 1.0% are black, and 0.8% are American Indian.

analytic models, we include a binary variable indicating if the observation is one of the 6% with a missing ethnicity indicator.

Finally, subsets of our full sample can also be linked to three additional variables. First, interns in the most recent years of our data were required to take the WEST-B teacher credential test in math, reading, and writing. We observe WEST-B scores for 56% of our sample. Since interns in Washington can take the WEST-B as many times as necessary to receive a passing score in each subject, we use the scores from the first time each intern took the test. Second, two TTIs consisting of 58% of our sample provided the undergraduate GPAs of their graduates. Lastly, the cooperating teachers of 26% of observed interns can be linked to student-level test score data, which allows us to calculate out-of-sample value-added measures of teacher performance for these cooperating teachers. We describe our measures of teacher value added in Appendix A.

3.2. Description of our analytic sample

The full intern sample consists of 8080 interns who completed student teaching by 2009 and received a teaching credential and endorsements to teach in Washington K-12 public schools. Of the 8080 interns in the sample, 2406 do not appear in the OSPI data by 2011. We refer to these interns as "not hired", meaning that they were not hired into a public or private K-12 job during the time that the OSPI data was observed. Note that "not hired interns" may include interns who were hired into a school (or other) position outside of Washington State, or hired into a school position after the last year of our dataset (2011), as well as interns who do not pursue or did not receive any position in a public or private school. Later, we address the issue of right-censoring in our analytic models.

The 5674 "hired" interns are observed in three different employment outcomes: public school teacher, public school non-teacher (e.g. paraeducator), and private school teacher. Several interns transition between these outcomes during our years of data, as illustrated by Fig. 2. For example, 87 interns are first hired as non-teachers in public schools before transitioning to a public teaching role, while 159 interns begin in a public teaching role before transitioning to a non-teaching position. There are some interesting differences between interns who follow different career paths.[16] For example, interns who transition from teaching to non-teaching positions within public schools tend to be older than interns who transition from non-teaching to teaching positions. Interns endorsed in STEM areas are far more likely to transition from teaching to non-teaching positions than vice versa, while interns endorsed in elementary education are far more likely to transition from private schools to public schools than vice versa.

In some of our exploratory analyses, it is useful to define one unique employment outcome for each intern. We define this employment outcome as each intern's first
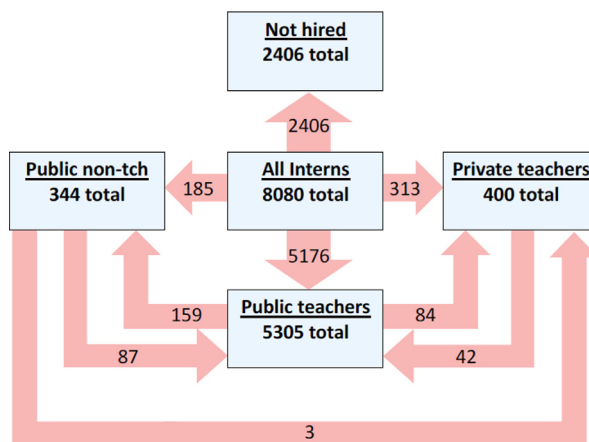

Fig. 2. Transitions of interns between observed hiring outcomes.

public school position after receiving his or her certification, or private school if we do not observe the intern employed in public schools. By these definitions, 271 interns (4.7%) are employed only in private K-12 teaching positions, while 185 (3.3%) were initially hired into public, non-teaching positions. The remaining 5218—64.5% of the 8080 interns in the sample—were hired into public, K-12 teaching positions, a proportion that is broadly consistent with what has been found using nationally representative data.[17]

3.3. Characteristics of cooperating teachers and internship schools

Assignment of interns to internship schools and cooperating teachers is determined by both state code and contractual arrangements between teacher training institutions and school districts. Washington state law requires the cooperating teacher to be highly qualified and to have a minimum of 3 years of full-time teaching experience. The state code also states that "field experiences provide opportunity to work in communities with populations dissimilar to the background of the candidate", which is often interpreted as a mandate to place interns in diverse internship schools.[18] Field placement agreements, on the other hand, are written contracts between TTIs and school districts that place interns in student teaching positions and match each intern with a cooperating teacher. These agreements usually allow interns to request schools in which to complete their internships, but these requests are circumscribed by the state's requirement to place students in diverse schools, as well as the needs of the internship building and the availability of cooperating teachers in relevant endorsement areas.

---

[16] Full descriptive statistics are available from the authors upon request.

[17] Ingersoll (2003) finds about 58 percent of new recipients of teaching credentials get a public teaching job within 4 years.

[18] The state code is from WAC 181-78A-264(3)(b)(ii), while the interpretation is from Jennifer McCleery of Western Washington University (personal communication, February 2014).

A recent report by The New Teacher Project (Greenberg et al., 2013) has criticized TTIs for not being more purposeful in their assignment of interns to internship schools and cooperating teachers, noting that "we are still looking for a program ... that includes an intensive screening of nominated cooperating teachers." Interestingly, the cooperating teachers in our sample were no more experienced in the year they supervised an internship, on average, than other teachers in the state in the same year, although they are more likely to be female and hold a masters degree. But we do see, for the sub-sample of cooperating teachers for whom it is possible to estimate value added, that cooperating teachers have marginally higher estimates of effectiveness in both reading (3.6% of a standard deviation of student performance, $p = .017$) and math (3% of a standard deviation, $p = .012$) than other teachers in the state.[19] This is preliminary evidence that the TTIs in our study are purposeful in selecting cooperating teachers to supervise internships.

Interns completed their student teaching in 1162 different schools across the state, and there is considerable variability in the characteristics of these internship schools, both within and across our participating institutions. Seemingly in contrast to the state mandate that internships occur in diverse schools, student teaching tends to take place in schools with fewer disadvantaged students than the average school in the state.[20] This may be because the majority of internships (50.5%) occur within 25 miles of the intern's training institution, and the average school within 25 miles of one of our participating institutions has fewer disadvantaged students (e.g., 30.3% FRL students) than the average school not within 25 miles of one of our participating institutions (e.g., 43.4% FRL students).

Since internships are not assigned randomly, we are cautious about drawing causal inferences about the relationship between these internship characteristics and the probability of employment. However, we estimate models that predict the level of advantage of an intern's schools (e.g., the percent of FRL students) as a function of intern qualifications (e.g., GPA or credential scores) and find little evidence that interns with strong qualifications are more likely to be assigned to more (or less) advantaged schools.

3.4. Descriptive picture of interns by labor market outcome

Our primary goal is to identify the teacher training experiences that are correlated with intern entry into the public teaching workforce. However, as we outline above, interns in our sample who are not employed as public school teachers may have been hired into non-teaching positions in public schools, as teachers in private schools, in non-schooling positions in Washington state, or into positions outside of Washington state (teaching or otherwise). As noted above, we cannot distinguish

between interns who are hired out-of-state and interns hired into the state into non-teaching positions (or who are unemployed), but we do know if prospective teachers are employed in private schools or in non-teaching positions in public schools. So, while our primary analysis focus exclusively on the likelihood of becoming a public school teacher (in Washington), we first explore whether interns end up employed in different positions in public schools or in private schools.

Table 1 compares interns by labor market outcomes along three dimensions: individual intern characteristics; characteristics of the intern's cooperating teacher; and characteristics of the intern's internship school. Interns hired into public or private teaching roles tend to be younger than those hired into non-teaching roles or who are not employed in the K-12 public teaching workforce in Washington. There is a large gender discrepancy between interns hired to teach in public versus private schools, and significant difference across endorsement areas, which is not surprising since private schools are not required to staff classes according to teacher endorsements.

There are relatively few differences across groups in terms of the characteristics of the cooperating teachers or internship schools. Interestingly, we do see that interns in schools with more advantaged students (as measured by percent minority students, percent FRL students, and state passing rates) are more likely to be hired as public school teachers than not hired interns. However, interns in schools with *more* teacher turnover are also more likely to be hired into public school positions. This is true in terms of the average stay ratio, but also in terms of the number of new teachers the internship school hires the following year. This points to the potential importance of interns being hired into the same school where they did their student teaching, an issue we return to in Section 5.

To further explore the factors that may be correlated with interns being employed in different types of positions, we restrict our sample to only those who completed their student teaching in 2003 or later (since we only have private school data beginning in 2004), and estimate multinomial logit models predicting which of the four labor market outcomes an intern experiences. Table 2 presents the estimated marginal effects from these models, where the reference group is not hired individuals. These estimates control for institution and internship year, as well as a host of other variables that we omit from the table for parsimony. See the notes of Table 4 for a full list of control variables.

The first column of Table 2 ("Full sample") contains estimates from the multinomial logit based on all observed individuals who completed their internships between 2003 and 2010. The marginal coefficients reported in this table are relative to the reference group and can be interpreted as the change in the probability of entering the observed group relative to not being hired, all else equal. For instance, interns endorsed in a STEM field are 3.1 percentage points less likely to become public school non-teachers and 23 percentage points more likely to be hired as public school teachers. Interns are also more likely to be hired into public school non-teaching jobs if they are older, male, and did not share the same endorsement as

---

[19] See Appendix A for details about the value-added estimates.
[20] For example, the average %FRL of internship schools in our sample is 34.4%, compared to 39.9% across the state ($p < .01$).

Table 1
Intern, cooperating teacher, and internship school characteristics by outcome.

	Public teaching role	Private teaching role	Public non-teaching role	Not observed hired
Full sample (N = 8080)	N = 5218	N = 271	N = 185	N = 2406
Intern characteristics				
Age	27.96** (7.66)	27.60* (7.53)	34.05 (9.70)	29.06 (9.01)
Male	23.78%	14.02%	34.05%**	22.98%
Non-white	8.80%	9.50%	8.11%	9.80%
Intern endorsement area				
STEM	13.97%**	6.27%	2.16%**	8.40%
Special education	13.11%**	.37%**	6.49%	6.57%
ELL	5.39%	4.06%	5.41%	4.36%
Elementary	63.53%**	81.55%**	57.84%**	67.87%
Other	37.01%*	21.03%**	48.11%**	34.50%
Cooperating teacher characteristics				
Age	44.95* (9.70)	45.10 (1.05)	44.30 (9.77)	45.49 (9.81)
Experience	15.02 (8.59)	15.16 (9.13)	14.54 (8.23)	15.15 (8.74)
Number prior observed interns	.38** (.93)	.66* (1.17)	.46 (1.01)	.51 (1.12)
Male	23.34%	13.28%**	32.43%**	22.32%
Master's degree	6.94%*	6.15%	56.22%	63.42%
Gender match	71.90%	81.18%**	73.51%	73.65%
Endorsement match	77.27%	75.65%	75.14%	79.14%
White/non-white match	82.8%	83.4%	84.9%	84.0%
Internship school characteristics				
Percent minority students	21.04** (17.46)	21.07 (15.49)	22.03 (16.47)	22.74 (17.88)
Percent FRL students	34.35** (2.61)	35.54 (19.72)	37.67 (19.44)	37.91 (2.64)
Standardized avg. passing rate	.28** (.83)	.32 (.85)	.26 (.81)	.22 (.84)
Standardized stay ratio	−.20** (.60)	−.20* (.67)	−.14 (.71)	−.12 (.66)
Number prior observed interns	7.68** (12.13)	11.42* (15.75)	1.74 (18.39)	9.34 (13.53)
Number new teachers hired next year	1.21** (1.54)	.88 (1.24)	1.01 (1.31)	.93 (1.29)
WEST-B sample (N = 4575)	N = 2837	N = 173	N = 107	N = 1458
Avg. WEST-B Score	272.14** (11.68)	272.18 (11.14)	267.03** (11.98)	270.75 (11.68)
VAM sample (N = 2083)	N = 1290	N = 80	N = 47	N = 666
Cooperating teacher VAM	.05 (.18)	.04 (.17)	.04 (.21)	.04 (.17)
GPA sample (N = 4535)	N = 2983	N = 145	N = 105	N = 1302
Undergraduate GPA	3.21 (1.06)	3.46* (.66)	3.29 (.73)	3.24 (1.02)

Significance levels for two-sided t-test relative to last column. Standard deviations in parenthesis.

* $p < .05$.
** $p < .01$.

their cooperating teacher. On the other hand, younger interns as well as those without a special education endorsement are more likely to be found in private schools.[21] The importance of endorsements is clear when examining individuals employed as teachers in public schools. STEM and ELL endorsed interns are more likely to be hired as public school teachers than teachers with an elementary endorsement, and younger interns are also more likely to be hired as public school teachers. Individuals serving their internships at schools with high stay ratios are also less likely to be a public school teacher, a result we examine in the next section.

The remaining columns of Table 2 report estimates from models adding covariates that are available for only a subset of interns—WEST-B score (averaged across math, science, and writing), undergraduate GPA, and cooperating teacher out-of-sample VAM—and are estimated only for the subset of interns for whom we have the appropriate data. The only notable additional finding from these models is that, all else equal, the probability of being hired into a public non-teaching position (relative to a public teaching position) decreases as the intern's WEST-B score increases while the probability of being hired into a public teaching job increases in the WEST-B score.

From this point on we restrict our focus to characteristics predictive of entry into the public teaching workforce by dropping the small number of interns hired into non-teaching or private school positions. One challenge in assessing the connection between training experiences and the labor market is the considerable heterogeneity we observe in the time between when interns complete their internship and when they are observed to be employed (see Fig. 3). In the next section we discuss the use of split population models to address this challenge, as well as a secondary analysis exploring the factors predicting whether interns are hired into the school in which their internship occurred.

---

[21] In the case of special education endorsements, there is only one intern in this sample endorsed in special education who was hired by a private school. Because of this, that coefficient is not identified nor reported in Table 3.

Table 2
Multinomial marginal effect estimates, reference group interns not hired (2003–2010).

	Full sample (N = 5749)			WEST-B sample (N = 4557)			GPA sample (N = 2610)			Coop VAM sample (N = 1599)		
	Public, non-teach	Private teacher	Public teacher	Public, non-teach	Private teacher	Public teacher	Public, non-teach	Private teacher	Public teacher	Public, non-teach	Private teacher	Public teacher
Intern age/10	.006*** (.002)	−.008** (.003)	−.050*** (.007)	.008*** (.002)	−.009** (.004)	−.050*** (.008)	.007** (.003)	−.001** (.006)	−.080*** (.010)	.006 (.004)	−.004 (.006)	−.050*** (.010)
Intern male (ref female)	.012*** (.005)	−.004 (.008)	.013 (.016)	.017*** (.006)	−.004 (.008)	−.007 (.018)	.014* (.007)	.009 (.012)	−.011 (.025)	.020** (.009)	−.004 (.014)	.063* (.034)
Intern non-white	−.001 (.011)	−.016 (.014)	.052 (.033)	−.001 (.011)	−.009 (.016)	.002 (.037)				.025 (.027)	−.005 (.029)	.0004 (.064)
Intern endorsed in STEM (ref elem)	−.031** (.014)	−.015 (.012)	.230*** (.027)	−.021 (.013)	−.015 (.013)	.212*** (.031)				−.010 (.020)	.040* (.021)	.210*** (.065)
Intern endorsed in ELL (ref not ELL)	.003 (.008)	−.007 (.012)	.122*** (.027)	.007 (.008)	−.002 (.012)	.112*** (.029)	.001 (.013)	−.012 (.026)	.140*** (.050)	−.008 (.019)	−.002 (.023)	.085 (.052)
Intern avg. WEST-B × 10				−.007*** (.002)	.001 (.020)	.010* (.006)						
Intern undergraduate GPA							−.001 (.003)	−.002 (.006)	−.004 (.012)			
Int school percent FRL students × 10	.002 (.001)	−.002 (.002)	.001 (.004)	.000 (.002)	−.002 (.002)	.001 (.005)	.003 (.002)	.000 (.003)	−.001 (.007)	.005 (.003)	.001 (.004)	.002 (.009)
Int school avg. passing rate (std)	.003 (.003)	.004 (.004)	.002 (.011)	.002 (.004)	.003 (.005)	−.0004 (.012)	.007 (.005)	.009 (.007)	−.011 (.016)	.002 (.007)	.005 (.009)	−.006 (.021)
Int school stay ratio (std)	.001 (.003)	.004 (.004)	−.034*** (.010)	.00008 (.003)	.005 (.004)	−.030*** (.011)	−.002 (.005)	.009 (.006)	−.032** (.015)	.007 (.006)	−.0005 (.008)	−.041** (.018)
Int school in city (ref suburb)	−.008* (.004)	−.006 (.005)	−.028* (.013)	−.011** (.005)	−.005 (.006)	−.019 (.016)	−.005 (.006)	−.00002 (.009)	−.020 (.021)	−.019* (.009)	.001 (.012)	−.008 (.027)
Int school in town (ref suburb)	−.001 (.007)	−.020 (.012)	−.022 (.025)	−.0006 (.007)	−.022 (.014)	.002 (.028)	.005 (.010)	−.015 (.018)	−.084** (.035)	−.010 (.016)	−.025 (.031)	.074 (.052)
Int school rural (ref suburb)	−.013* (.007)	−.015 (.010)	−.029 (.022)	−.008 (.008)	−.026** (.013)	−.015 (.025)	−.006 (.011)	−.018 (.016)	−.036 (.033)	−.009 (.014)	.006 (.017)	−.047 (.041)
Coop tch experience × 10	−.002 (.003)	−.001 (.004)	.016 (.010)	−.003 (.004)	−.001 (.005)	.010 (.010)	−.002 (.005)	.004 (.007)	.002 (.010)	.000 (.006)	−.002 (.009)	.020 (.020)
Coop tch gender match	.007 (.005)	.006 (.008)	.002 (.016)	.011* (.005)	.003 (.008)	.0001 (.018)	.012* (.007)	.016 (.012)	−.023 (.023)	.002 (.009)	.007 (.014)	.054 (.033)
Coop tch endorsement match	−.015*** (.005)	−.007 (.007)	.015 (.018)	−.012** (.006)	−.006 (.008)	.002 (.021)	−.006 (.008)	.005 (.012)	.002 (.026)	−.014 (.010)	−.019 (.014)	.110*** (.034)
Coop tch avg. VAM										.022 (.022)	−.032 (.031)	.090 (.069)

Samples include hired interns who did their student teaching in 2002 or later. See Table 4 for other notes.
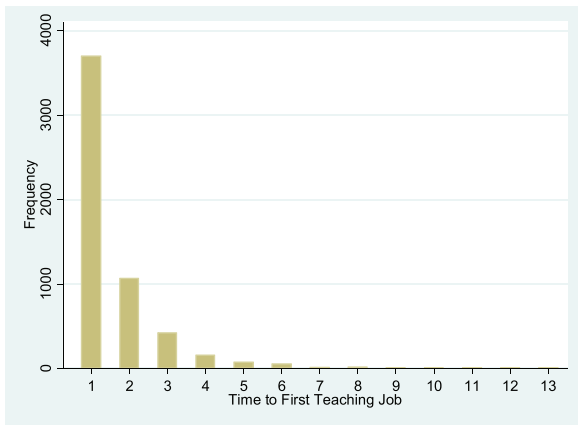
Fig. 3. Time to first teaching job for interns hired into teaching job.

4. Analytic approach

4.1. Split population model

To assess the relationship between internship experiences and employment as a public school teacher, it is typically assumed that the probability of employment for individual i depends on a latent variable, Y_i^*, and the observed outcome depends on whether this latent variable exceeds some threshold, c, that determines the hiring decision:

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* > c, \text{ intern } i \text{ is employed in a} \\ & \text{teaching position} \\ 0 & \text{if } Y_i^* \leq c, \text{ intern } i \text{ is not employed in a} \\ & \textit{teaching position} \end{cases} \quad (1)$$

A common econometric approach is to formulate (1) as a binary choice model and estimate the marginal effect of explanatory variables **X** on the probability of observing $Y_i = 1$. However, this approach ignores three related aspects of the transition from student-internship into the labor market. First, as demonstrated by Fig. 3, there is considerable heterogeneity in the time it takes an individual to be hired into a teaching job. Binary choice models produce no information regarding the time it takes to be hired; they simply model whether hiring occurs or not. But it is conceivable that characteristics of an internship experience differentially impact the likelihood of being hired and the timing of that hire. Binary choice models confound these impacts and tell us nothing of the timing of hire.

Second, our data are right-censored. Specifically, there are likely to be a considerable number of interns, especially those completing their internships late in our sample, who will successfully find a public teaching job after the last year they are observed in our dataset. The standard approach in this setting is to use survival analysis to model the time until each intern is hired into the workforce. But survival analysis assumes that all interns will eventually be hired into the workforce, which brings us to the third issue: many interns never become teachers and never would

become teachers even in the absence of censoring. This subpopulation of interns may, or may not, differ in measurable ways from those who search for and do not find employment. To account for the potential differential impacts of observable characteristics on hiring and the timing of hiring, the right-censored data, and the fact that a subset of interns will never find employment, we employ a split-population model.[22]

Split-population models simultaneously estimate the impact of covariates on the timing and probability of an event. Specifically, split population models explicitly account for the possibility that some individuals have a hazard of zero; i.e. those interns who will never have a teaching job, either because they choose not to pursue a job or because they will never be hired. Split population models are popularly used to explore the reoccurrence of cancers[23] and have been used by economists to study job placement and timing (Kyyrä & Ollikainen, 2008; Swaim & Podgursky, 1994), criminal recidivism (Schmidt & Witte, 1989), survival of financial institutions (DeYoung, 2003; Maggiolini & Mistrulli, 2005), and smoking cessation (Douglas & Hariharan, 1994).

As noted in Swaim and Podgursky (1994), a split-population formulation of job placement is stylized in that it assumes that interns make a one-time decision whether or not to pursue a teaching position. This is unrealistic in that it rules out intentional delays to entering the teacher workforce, but as Swaim and Podgursky note, a single-population survival analysis approach makes the even less realistic assumption that all interns who complete student teaching decide to pursue and will ultimately receive a teaching job.

In the split-population framework, we define the latent variable Y_i^* as an indicator of whether intern i will eventually be hired into a teaching job, and define T_i^* as the number of years from an intern's student teaching experience to his or her placement in a public K-12 teaching job. T_i^* is defined only for interns who are eventually hired ($Y_i^* = 1$). T_i^* is assumed to have a distribution function $f(t, \mathbf{Z}_i)$ where $\mathbf{Z}_i$ is a vector of observable characteristics for intern i. Define $F(t, \mathbf{Z}_i) = Pr(T_i^* \leq t), t > 0$ as the corresponding cumulative distribution. Note that because of right-censoring, we do not observe T_i^* and Y_i^* for all the interns in our sample who will eventually be hired. Thus, define T_i as the time to first job for interns who are observed to be hired ($Y_i = 1$) and the time to censoring for interns who are not ($Y_i = 0$). The goal of this part of our analysis is to use our observations of T_i and Y_i for each intern in our sample to make inferences about the factors that influence T_i^* and Y_i^*.

---

[22] We experiment with both logit and hazard models and find that the primary findings from these models are consistent with the estimates from the split population model. In fact, as we increase the number of years considered in the logit model—e.g., hiring after 1 year, hiring after 3 years, hiring after 5 years, etc.—the logit estimates get closer and closer to the "hire" estimates from the split population model.

[23] Split population models are called "cure models" in the medical literature because they assume that a subset of individuals are "cured" and will never have a reoccurrence of cancer, for example.

We consider a model for T_i^* and Y_i^* that splits our observations into two groups of interns, one of which will eventually be hired and the other of which will not.[24] The conditional distribution functions for T_i^* are defined as:

$$f(t_i|Y_i^* = 1, Z_i) = Pr(T_i^* = t_i|Y_i^* = 1, Z_i) = g(t, Z_i) \qquad (2)$$

$$F(t_i|Y_i^* = 1, Z_i) = Pr(T_i^* < t_i|Y_i^* = 1, Z_i) = G(t, Z_i) \qquad (3)$$

Let $\delta_i = Pr(Y_i^* = 1|Z_i)$. For interns who are hired during the sample period, we observe $Y_i = Y_i^* = 1$ and $T_i = T_i^* = t_i$. Thus can write joint distribution of the observed data for these interns as:

$$Pr(Y_i = 1, T_i = t_i|Z_i) = Pr(Y_i^* = 1|Z_i)Pr(T_i^* = t_i|Y_i^*$$
$$= 1, Z_i) = \delta_i g(t_i, Z_i) \qquad (4)$$

In contrast, the interns who are not hired during the sample period ($Y_i = 0$) might never be hired ($Y_i^* = 0$) or might be hired after the sample period ($Y_i^* = 1$ and $T_i^* > t_i$). The joint distribution of the observed data for interns with $Y_i = 0$ is:

$$Pr(Y_i = 0, T_i = t_i|Z_i) = Pr(Y_i^* = 0|Z_i)$$
$$+ Pr(Y_i^* = 1|Z_i)Pr(T_i^* > t_i|Y_i^* = 1, Z_i) \qquad (5)$$
$$= (1 - \delta_i) + \delta_i(1 - G(t_i, Z_i))$$

Combining (4) and (5) and assuming independence across observations yields the likelihood function for the observed data Y_i and T_i[25]:

$$L = \prod_{i=1}^{n}[\delta_i g(t_i, \mathbf{Z}_i)]^{y_i}[1 - \delta_i + \delta_i(1 - G(t_i, \mathbf{Z}_i))]^{1-y_i} \qquad (6)$$

Within this likelihood, we can specify a functional form for both δ_i and $G()$ and estimate coefficients relating the observed characteristics of each intern to the probability of getting hired (δ_i) and the time to hire ($G(t_i, \mathbf{Z}_i)$). The split-population literature provides a number of options. For the results presented below, we model δ_i as a logit in $\mathbf{Z}$:

$$\log\left(\frac{\delta_i}{1 - \delta_i}\right) = \gamma\mathbf{Z}_i + \varepsilon_i$$
$$\Rightarrow \delta_i = \frac{\exp(\gamma\mathbf{Z}_i + \varepsilon_i)}{1 + \exp(\gamma\mathbf{Z}_i + \varepsilon_i)} \qquad (7)$$

In (7), γ is a vector of coefficients representing the correlation between each observable intern characteristic and the log odds of the intern eventually being hired.

Our primary results use an exponential model for the time to hire function $G()$:

$$G(t_i, \mathbf{Z}_i) = 1 - \exp\left(-\frac{t_i}{\exp(\beta\mathbf{Z}_i)}\right) \qquad (8)$$

In (8), β is a vector of coefficients representing the correlation between each observable intern characteristic

and the slope of the hazard curve representing time-to-hire.

One drawback of the split population formulation is that it is impossible to know from the observed data whether an unemployed intern will never be hired or is simply right-censored. Split population models use the functional form of the hazard function to help distinguish between these two possible outcomes; that is, using data on probability of hire and time-to-hire for interns observed for many years, the model essentially imputes these values for interns observed for few years who are not observed to be hired. As pointed out by Jaggia (2011), it is possible that eventually all interns would be hired and that the split population model incorrectly identifies some of these as being cured (i.e., never to be hired). However, this is more likely to occur for datasets containing relatively few observed periods, while our dataset leaves ample time for the earliest interns to be hired. Yet, even for our first cohort of interns who completed their internship in 1998, 22.1% are not observed hired by 2011. It is very unlikely that these individuals are continuing to search for a teaching job, so we believe that a split population model is a more accurate representation of the reality that these interns *will never be hired* than a hazard model that assumes that the data are just right-censored.[26]

For completeness, we follow Jaggia's advice to estimate split population models and duration models with various specifications of $G()$ and to determine if the results are robust to the choice of fail density.[27] Here we simply note that for various specifications of both split population and hazard models, we find qualitatively similar results to what we later present.[28]

We discuss our estimates of γ and β in the next section, but these coefficients can be difficult to interpret because they describe related dimensions of the same outcome: the former describes the probability of eventually getting a public teaching job, while the latter describes the time until the intern is hired. To ease the interpretation of our results, we calculate marginal effects for each covariate 1 and 5 years after the student completes his or her internship.[29] These marginal effects can be interpreted as

---

[24] We assign $Y = 1$ to individuals finding a job which, in traditional split population terminology, are "failures."

[25] We maximize this likelihood using the user-written STATA module CUREREGR (Buxton, 2007). We do not cluster standard errors at the institution level because we are not interested in any institution-level covariates.

[26] Another shortcoming of split population models is that omitted variables in the "probability of hire" equation (7) are assumed to be independent of omitted variables in the "time to hire" equation (8). This is not a particularly realistic assumption, as unobserved characteristics that make interns less likely to be hired are also likely to make interns less likely to be hired quickly if they are hired. With that said, a hazard model assumes that there is no error in the "probability of hire" model (because every intern is assumed to be hired), while a logit model assumes that there is no error in the "time to hire" model (because the outcome is binary). Under these circumstances, the split population model makes the least problematic implicit assumptions.

[27] Specifically, when we allow the time to hire to have both a shape and scale parameter using a Weibull or a Gamma distribution, the estimates for the scale parameter are very similar to the estimates for the scale parameter using the exponential distribution while the estimates for the shape parameter are not statistically significant.

[28] These alternative models are available upon request.

[29] The marginal effects are estimated by calculating the average change in the probability of being hired across all observations given a unit change in the independent variable.

Table 3
Split population estimates for hiring as public school teacher vs. not observed hired (1998–2010).

	Full sample (N = 7624)			WEST-B sample (N = 4295)			GPA sample (N = 4433)			Coop VAM sample (N = 1956)		
	Hired (SD)	Time (SD)	1 year ME / 5 year ME	Hired (SD)	Time (SD)	1 year ME / 5 year ME	Hired (SD)	Time (SD)	1 year ME / 5 year ME	Hired (SD)	Time (SD)	1 year ME / 5 year ME
Intern age/10	−.299*** (.050)	.050* (.023)	−.025 / −.040	−.443*** (.111)	.052 (.042)	−.028 / −.052	−.329*** (.052)	.066* (.030)	−.025 / −.033	−.140 (.096)	−.009 (.046)	−.006 / −.018
Intern male (ref female)	−.011 (.105)	.029 (.042)	−.007 / −.004	−.122 (.190)	−.015 (.063)	−.002 / −.008	−.036 (.108)	.020 (.055)	−.007 / −.007	.077 (.210)	−.072 (.100)	.021 / .021
Intern non-white	−.484*** (.135)	−.070 (.057)	−.005 / −.049	−.550* (.223)	−.220** (.081)	.013 / −.012	−.586*** (.150)	.060 (.089)	−.034 / −.057	−.688 (.461)	−.123 (.207)	−.021 / −.091
Intern endorsed in STEM (ref elem)	.924*** (.181)	−.281*** (.062)	.108 / .106	.701* (.291)	−.343*** (.092)	.109 / .124	.972*** (.191)	−.327*** (.086)	.115 / .093	1.513*** (.469)	−.039 (.108)	.072 / .148
Intern endorsed in SPED (ref elem)	.974** (.297)	−.246* (.096)	.103 / .114	.655 (.445)	−.464** (.135)	.144 / .141	.839** (.298)	−.202 (.117)	.075 / .075	.743** (.278)	−.401*** (.106)	.144 / .135
Intern endorsed in ELL (ref not ELL)	.700** (.245)	−.069 (.073)	.047 / .075	.533 (.438)	−.188 (.111)	.063 / .083	.510* (.242)	−.046 (.101)	.019 / .033	.447 (.436)	.140 (.191)	−.010 / .029
Intern avg. WEST-B				−.006 (.076)	.049* (.025)	.001 / .001						
Intern undergraduate GPA							.040 (.042)	.015 (.022)	−.003 / −.002			
Int school percent FRL students	−.046 (.035)	−.003 (.014)	.000 / .000	.002 (.090)	.006 (.024)	.000 / −.001	−.058 (.036)	−.017 (.019)	.001 / −.001	−.002 (.006)	.004 (.003)	−.001 / −.001
Int school avg. passing rate (std)	−.105 (.076)	−.023 (.031)	.000 / −.006	−.065 (.218)	−.037 (.059)	.006 / .003	−.154* (.077)	−.030 (.038)	.000 / −.005	.058 (.138)	.071 (.066)	−.013 / −.004
Int school stay ratio (std)	−.199** (.068)	.034 (.029)	−.017 / −.028	−.147 (.152)	.084 (.049)	−.024 / −.033	−.194** (.072)	−.002 (.037)	−.006 / −.013	−.205 (.123)	.060 (.062)	−.025 / −.038
Int school in city (ref suburb)	−.232* (.102)	.047 (.039)	−.020 / −.031	−.361 (.210)	.021 (.063)	−.018 / −.037	−.170 (.106)	.025 (.051)	−.008 / −.014	−.086 (.188)	−.113 (.086)	.020 / .005
Int school in town (ref suburb)	−.445** (.145)	−.056 (.067)	−.009 / −.038	−.663* (.262)	−.182 (.101)	.011 / −.034	−.504** (.148)	−.009 (.083)	−.014 / −.030	−.340 (.289)	−.095 (.148)	−.001 / −.037
Int school rural (ref suburb)	−.394** (.146)	.063 (.068)	−.033 / −.049	−.533 (.326)	.036 (.110)	−.029 / −.059	−.423* (.138)	.046 (.074)	−.034 / −.042	−.590* (.259)	−.001 (.145)	−.038 / −.089
Coop tch number prior interns	−.142** (.046)	−.047* (.022)	.004 / −.010	−.120 (.087)	−.045 (.032)	.009 / .002	−.143** (.050)	−.025 (.027)	−.002 / −.015	−.117 (.094)	−.046 (.046)	.003 / −.010
Coop tch gender match	−.048 (.106)	.030 (.042)	−.010 / −.010	−.170 (.194)	−.060 (.063)	.007 / −.001	.007 (.100)	.059 (.050)	−.016 / −.011	.079 (.211)	−.018 (.099)	.009 / .014
Coop tch endorsement match	.046 (.112)	−.019 (.047)	.009 / .011	−.053 (.246)	.015 (.082)	−.004 / −.008	.164 (.109)	−.001 (.055)	.002 / .008	.495* (.217)	−.012 (.102)	.033 / .074
Coop tch avg. VAM										.514 (.450)	−.127 (.213)	.058 / .079

Samples include all interns hired as public school teachers or not observed hired into any position. See Table 4 for other notes.

the expected change in the probability of being hired 1 or 5 years after completing the internship for each unit change in the covariate.

4.2. Hiring into internship school

One intriguing finding from our exploratory analysis is that 806 of the 5218 interns hired into public schools (15.4%) were hired by the school where they did their student teaching, suggesting that student teaching may serve not only training purposes, but also provides schools with information about the ability and fit of prospective teachers. We employ a logit model to explore the probability, θ_i, that an intern is hired into his or her internship school:

$$\log\left(\frac{\theta_i}{1 - \theta_i}\right) = \lambda \mathbf{Z}_i + \varepsilon_i \tag{9}$$

We first estimate this model for all hired teachers, so the dependent variable (in Eq. (9)) is the log odds of being hired at one's internship school, relative to being hired at another school. This, however, ignores the fact that internships may have occurred in schools that did not have any available openings when interns were seeking employment. Given this, we also estimate this model for the subset of interns who did their student teaching at a school that hired at least one new teacher the following year, and further control for the number of interns the school hired. We transform all logit coefficients to marginal effects (calculated at the intern level) to ease in interpretation of our results.

5. Results

5.1. Probability and timing of hiring as public school teacher

Table 3 reports the selected estimated coefficients and marginal effects from four specifications of the split population model. For each model, we report the vector of estimated coefficients $\hat{\gamma}$ in the "Hired" column (these coefficients are on the log odds scale). We stress that these coefficients should not necessarily be interpreted as reflecting the hiring preferences of employers or employees. Positive values of these coefficients represent a positive correlation between the variable and the probability of eventually entering the teaching workforce. We also report the vector of estimated coefficients $\hat{\beta}$ in the "Time" column, which represent the relationship between each variable and the time-to-hire. The "Hired" and "Time" coefficients can be difficult to interpret together: the first represents the probability of *eventual* hire, while the second determines the slope of the hazard curve for hiring for interns who eventually be hired. Because of this, we also report marginal effects for each coefficient for the probability of hiring 1 and 5 years after student teaching. To further solidify intuition, we plot fitted probabilities of hire over time for selected covariates in Fig. 4. In these plots, the vertical distance between the curves at each time point corresponds to the marginal effect at that time.

The first set of results in Table 3 reports selected coefficients from a split population model estimated for the full sample of interns (columns 1–3). The full list of control variables is noted at the bottom of Table 4; all models control for an intern's training institution, internship year, and internship term. This is important because we observe large disparities in placement rates between participating institutions and internship years.[30]

Several intern characteristics are correlated with the probability and timing of an intern being employed in a public teaching job. All else equal, younger interns are more likely to be in a public teaching job: an increase of 10 years of age is associated with a 2.5 percentage point decrease in the probability of being employed in a public school after 1 year, and is correlated with a 4.0 percentage point decrease in the probability of being hired into a public school after 5 years. These marginal effects—the vertical distance between the time-to-hire curves for interns of average age (28 years old) and 10 years older than average age (38 years old)—can be seen increasing over time in Fig. 4a. One possible explanation for the age finding is that school systems prefer to hire younger interns believing in the traditional practice of hiring recent college graduates who can dedicate an entire career to teaching (Hess, 2009). But, it is also possible that older interns are career changers who may be less likely to seek a teaching job, even having obtained a teaching credential.[31]

Although the raw difference in observed employment rates for white and non-white interns is not statistically significant (72.5% for white vs. 70.0% for non-white, $p = .130$), the split population estimates in Table 4 suggest that non-white interns are significantly less likely to be employed (.5 percentage points after 1 year and 4.9 percentage points after 5 years), all else equal, than white interns (these differences are also plotted over time in Fig. 4b).[32] This seemingly runs contrary to the rhetoric about the desirability of diversifying the teacher workforce and existing empirical evidence (Boyd, Lankford, Loeb, Ronfeldt, & Wyckoff, 2011). To dig deeper into this finding we estimate models that interact the non-white indicator with indicators for each institution to assess whether it is consistent across training programs. In these specifications each interaction term (and the main effect)

---

[30] For example, not surprisingly given the economic downturn, our estimates suggest that there was a sharp drop in the probability of getting hired for interns who graduated in 2008 or later.

[31] Also, as we note in the data section, teachers who obtain a Masters degree (and we do not observe the type of degree for non-hired interns) tend to be older so the age result may also be picking up some of the supply or demand effects associated with the receipt of an MA versus a BA degree.

[32] A possible explanation for why non-white interns are less likely to find employment, despite little difference in average employment rates, is that they are more likely to do their student teaching in schools with disproportionately high teacher turnover. As we will discuss later, interns from schools with high teacher turnover are more likely to find employment, all else equal, and the average standardized internship school stay ratio is −0.23 for non-white interns and −0.17 for white interns ($p = 0.022$).
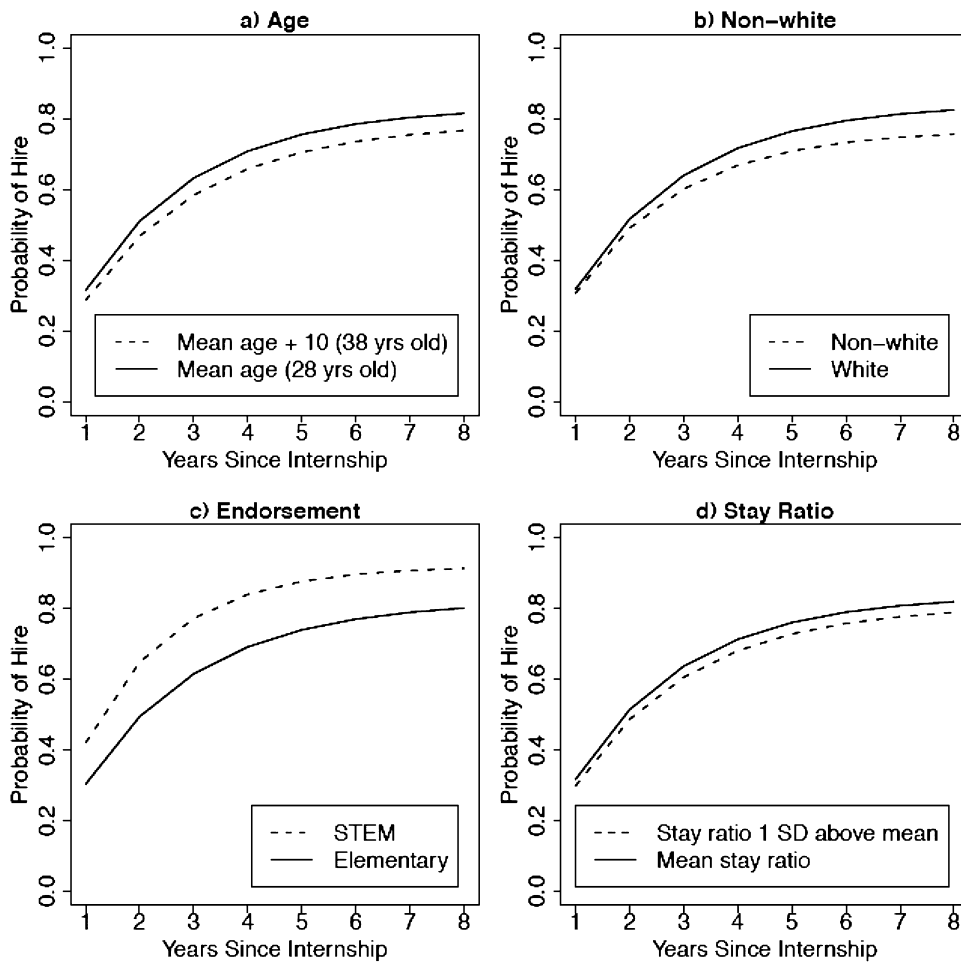
Fig. 4. Fitted probabilities of hire for selected variables (at the means).

is negative; that is, non-white interns are less likely to be employed, all else equal, regardless of the institution they attended.[33] We also interact the non-white indicator with indicators for internship school geographic location (west of Puget Sound area, Puget Sound area, western half of state, and other) and find that the interaction between non-white and the western half of state is significant and negative, though why employment prospects for prospective minority teachers ought to be diminished in the part of the state with higher minority student populations, particularly in the Puget Sound region, is not clear.

Finally, and perhaps most importantly, we estimate a number of specifications of the split population model that include separate identifiers for the race/ethnicity of the interns: American Indian, Asian, black, and Hispanic interns (with the reference category being white interns). Interns of each non-white ethnicity are less likely to be

employed than white students, all else equal, but only the coefficients for American Indian and Asian are statistically significant.[34] These findings are mostly robust to the inclusion of internship district-by-year fixed effects in the split population model (i.e., American Indian and Asian interns are still significantly less likely to be hired, all else equal), although the sign for Hispanic interns flips in this model.[35]

The bottom line is that the race/ethnicity results are a bit puzzling and difficult to interpret as it is not clear whether they are driven by the preferences of hiring officials or prospective employees, who might have differential employment opportunities outside of public schools, or outside of Washington. That said, these results

---

[33] There are some sizeable differences in the proportion of minority interns graduating from the six institutions in our sample—for example, 15.0% of interns from UW-Seattle are non-white, compared to only 7.6% of interns at Western Washington—but the findings on non-whites cannot be driven by differences in employment prospects associated with institution since the model includes training program fixed effects.

[34] The log odds coefficients for probability of eventual hire and corresponding standard errors for each category are −0.839 (SE = 0.38) for American Indian, −0.57 (SE = 0.19) for Asian, −0.05 (SE = 0.42) for black, and −0.33 (SE = 0.23) for Hispanic.

[35] The log odds coefficients from the model with internship district-by-year fixed effects for probability of eventual hire and corresponding standard errors for each category are −0.821 (SE = 0.38) for American Indian, −0.44 (SE = 0.18) for Asian, −0.29 (SE = 0.36) for black, and 0.18 (SE = 0.26) for Hispanic.

Table 4

Logit marginal effect estimates for hiring into internship school vs. hiring into other school (1998–2010).

Sample	Hired	Open	WEST-B	GPA	Int VAM	Coop VAM
Sample size	N = 5218	N = 2970	N = 1524	N = 1824	N = 727	N = 664
	ME (SD)	ME (SD)	ME (SD)	ME (SD)	ME (SD)	ME (SD)
Intern age/10	.001 (.007)	.012 (.010)	.017 (.015)	.000 (.014)	.025 (.020)	.052** (.019)
Intern male (ref female)	−.010 (.013)	−.022 (.019)	−.060* (.027)	−.023 (.024)	−.065 (.042)	−.053 (.041)
Intern non-white	.047** (.016)	.045 (.025)	.046 (.035)	.026 (.038)	.010 (.053)	−.079 (.064)
Intern endorsed in STEM (ref elem)	−.002 (.018)	−.001 (.027)	−.005 (.038)	−.039 (.038)	.025 (.064)	.011 (.063)
Intern endorsed in SPED (ref elem)	.009 (.028)	−.011 (.044)	−.022 (.062)	−.025 (.059)	−.102 (.131)	.030 (.105)
Intern endorsed in ELL (ref not ELL)	−.029 (.023)	−.030 (.035)	−.031 (.045)	−.021 (.047)	−.081 (.076)	−.144 (.096)
Intern avg. WEST-B × 10			.023* (.011)			
Intern undergraduate GPA				.000 (.010)		
Intern future VAM					.095 (.079)	
Int school percent FRL students × 10	−.002 (.004)	.000 (.006)	.001 (.009)	.002 (.009)	.004 (.012)	−.002 (.013)
Int school avg. passing rate (std)	−.012 (.009)	−.014 (.014)	−.014 (.021)	−.024 (.018)	.004 (.027)	−.012 (.029)
Int school stay ratio (std)	−.028** (.009)	−.020 (.015)	−.006 (.022)	−.043* (.020)	−.029 (.029)	−.017 (.030)
Int school in city (ref suburb)	−.025* (.012)	−.027 (.018)	−.010 (.026)	−.034 (.023)	−.019 (.037)	−.038 (.038)
Int school in town (ref suburb)	.041* (.019)	.081** (.030)	.126** (.042)	.087* (.037)	.205 (.058)	.138* (.059)
Int school rural (ref suburb)	.054** (.018)	.055 (.029)	.053 (.043)	.081* (.037)	.051 (.051)	.072 (.053)
Coop tch experience × 10	−.003 (.006)	−.003 (.010)	.003 (.013)	−.023 (.013)	−.018 (.020)	−.002 (.021)
Coop tch gender match	.002 (.013)	−.016 (.019)	−.019 (.027)	−.017 (.024)	−.042 (.042)	−.024 (.041)
Coop tch endorsement match	.022 (.015)	.045* (.022)	.048 (.033)	.004 (.029)	−.044 (.040)	.025 (.045)
Coop tch avg. VAM						.080 (.085)

Samples include all interns hired into public schools. Significance levels from two-sided *t*-test: *p < .05; **p < .01; ***p < .001. All models include indicators for internship year, training institution, and internship term, as well as intern gender, interactions between the number of multiple endorsements and teacher endorsement areas, indicators for intern prior and current school experience, and missing race indicator; internship school enrollment, percent minority students, indicators for Idaho/Oregon borders, and observed number of prior interns at the internship school; and indicators for cooperating teacher masters degree, male, and observed number of prior interns, as well as interactions between a supervisor's and intern's endorsement, race, and gender.

are consistent with findings in White et al. (2013) that non-white certificated individuals are less likely to enter the teaching workforce in Illinois. In each case, of course, it is also possible that the findings reflect some omitted variable that is correlated with both the non-white indicator and probability of employment. For example, each minority intern sub-group has significantly lower average WEST-B scores than white interns.[36] We return to this point below when we discuss the findings for models that control for measures of academic proficiency.

Not surprisingly given the evidence that school systems tend to report greater difficulty recruiting and retaining teachers with certain endorsements, we see strong evidence that a teacher's endorsement area predicts the probability of employment.[37] Relative to interns endorsed in elementary education, interns endorsed in STEM are far more likely to be employed all else equal (10.8 percentage points after 1 year and 10.6 percentage points after 5 years, shown relative to elementary in Fig. 4c). STEM and elementary endorsed teachers may not be a natural

comparison group as STEM is a secondary endorsement and the labor market for secondary teachers may differ from that of elementary teachers. Yet, the large advantage STEM holds over other elementary endorsement holds when compared to other secondary fields. Relative to all other secondary endorsements, STEM endorsed teachers are 6 percentage points more likely to be employed after 1 year and 12.1 percentage points more likely after 5 years.

Endorsement advantages are not limited just to STEM teachers. Special education endorsements have similarly large employment impacts with a 10.3 percentage point increase in employment probability after 1 year and 11.4 percentage points increase after 5. Interns with an endorsement to teach ELL classes are also more likely to be employed, all else equal, than interns without an ELL endorsement (4.7 percentage points after 1 year and 7.5 percentage points after 5 years).[38]

One might expect that cooperating teachers or internship schooling characteristics would influence the likelihood of workforce entry, either directly through the training that interns receive or because the reputation of a school or recommendation of the cooperating teacher would carry weight when interns sought a job. In particular, discussions with school hiring officials suggest it is common for cooperating teachers to write letters of recommendation for prospective teachers, few of these

---

[36] The average standardized WEST-B score is 0.028 for white interns, −0.093 for Asian interns, −0.276 for American Indian interns, −0.390 for black interns, and −0.426 for Hispanic interns.

[37] Since interns can hold an endorsement in more than one area, our model contains interactions between an indicator for whether an intern holds multiple endorsements and the STEM, special education, other, and elementary indicators. ELL, unlike the other categories, is a secondary endorsement, which means that interns endorsed in ELL *must* be endorsed in another area. We therefore do not interact the ELL and multiple endorsement variables. The STEM and special education coefficients are therefore interpreted relative to elementary education, while the ELL coefficients are measured relative to all interns not endorsed in ELL.

[38] When we explore models with interactions between endorsement areas and year of internship, only one of the 22 interactions is statistically significant at the 95 percent confidence level, which is about what we would expect by random chance. Thus we conclude that the impact of endorsement area on probability of hiring is consistent over the years in our sample.

internship variables are significant predictors of the probability and timing of workforce entry.[39] The only cooperating teacher characteristic that is a significant predictor of employment is the number of interns from participating institutions mentored by each cooperating teacher in prior years during the period of our data, which is negatively correlated with probability of employment. It is unclear whether this might be related to the nature of the training received by interns or the guidance they might receive from more experienced cooperating teachers. For instance, one could imagine that more experienced cooperating teachers are teachers that school systems feel need extra help in the classroom so they are assigned more interns, possibly affecting the quality of the training interns receive and hence their desirability as applicants. On the other hand, more experienced cooperating teachers may provide interns with different information about their prospects as teachers, affecting their supply decisions.

Just as Ronfeldt (2012) finds little correlation between the characteristics of the students in a teacher's internship school and workforce outcomes, we find little evidence that internship school student characteristics are predictive of hiring outcomes.[40] However, two other internship school characteristics do seem to matter for K-12 employment prospects. First, interns who complete their student teaching in cities, towns, and rural areas are less likely to be employed in a public teaching position than interns who did their student teaching in suburban areas, all else equal. Second, the probability of employment decreases as the average amount of teacher turnover in an intern's internship school decreases (i.e., as the school's stay ratio increases): a one standard deviation increase in the stay ratio is correlated with a 1.7 percentage point decrease in the probability of employment after 1 year and a 2.8 percentage point decrease after 5 years (these differences are shown over time in Fig. 4d). This finding is interesting in that it conflicts with Ronfeldt (2012), who finds that teachers who complete their student teaching in schools with low teacher turnover are both more effective (in terms of value added) and stay in teaching longer, attributes that should make interns more desirable job candidates.

It is possible that hiring officials are unaware of the connection between internship school and the outcomes of interns as teachers. It is also possible that Ronfeldt's findings on effectiveness and attrition are biased by sample selection. For example, if only the most motivated interns from schools with low teacher turnover enter the workforce, and these same interns are more effective as teachers and more likely to stay in the profession longer, then Ronfeldt's findings may be driven by the impact of student teaching on workforce entry, not the impact of the student teaching on effectiveness and retention. On the other hand, it could be that interns serving at a low-turnover school are more likely to apply for jobs at similar schools, thereby reducing their probability of being hired because they concentrate their search at schools making few new hires. We test this possibility by adding an interaction term to the hiring model between internship school stay ratio and whether the internship school hired at least one new teacher the following year. This interaction is not statistically significant, which suggests that something other than teacher turnover is driving this result.

Fortunately, another possibility exists: it is possible that schools use internships as screening devices for future hiring. If this is the case, then students completing internships at schools with a higher stay ratio would be less likely to be hired at their internship school because of its low teacher turnover and these individuals would be unable to demonstrate their effectiveness at a school that was about to hire a teacher. We explore this possibility in the next sub-section.

The split population model estimated for the full sample contains little in the way of controls for individual heterogeneity, and as we discussed in regards to the non-white findings, omission of these controls may bias the estimates from the full sample. With this in mind, the final three models in Table 4 report estimates from models that add covariates that are available for only a subset of interns—WEST-B score (averaged across math, science, and writing), undergraduate GPA, and cooperating teacher out-of-sample VAM—and are estimated only for the subset of interns for whom we have the these additional data elements. For each of these subsets we find little evidence that measures of intern academic proficiency, or the effectiveness of an intern's cooperating teacher, are correlated with the probability of hiring as a public school teacher.[41] Importantly, non-white interns are still less likely to be hired even in models that control for intern academic proficiency[42], although this does not rule out

---

[39] Similarly, one might have hypothesized that the training experience of interns would be enhanced by a race/ethnicity or gender match between cooperating teacher and intern or, perhaps most importantly, by being matched to a cooperating teacher with the same endorsements. There is evidence that matches between teacher and student demographics can influence teacher productivity and speculation that this may be related to teachers ability to connect with students given similar backgrounds/perspectives (Ehrenberg, Goldhaber, & Brewer, 1995), so it is not outlandish to imagine we would see these sort of effects with cooperating teachers and interns. But, as it turns out, this does not to be the case at least in terms of the probability of eventual K-12 public school employment.

[40] Note that we do not include the percent of underrepresented minority students in the internship school in these models because, across schools in Washington State, this variable is highly collinear with the percent of free/reduced lunch students in the school ($r = 0.67$).

[41] We also experiment with a split population model that includes internship district-by-year fixed effects, and find that most of our are qualitatively similar: the probability of hire decreases as age increases; interns endorses in STEM and special education are more likely to be hired than interns endorsed in elementary education; and interns endorsed in ELL are more likely to be hired than interns not endorsed in ELL. One important result that changes is the coefficient in internship school stay ratio; the probability of hire still increases as teacher turnover increases, but the coefficient is less than half as large and not statistically significant. This is not surprising given that much of the variation in the stay ratio is cross district (40%) rather than within.

[42] When we decompose the non-white indicator into individual ethnicity indicators, we find that (as in the full model) Asian and American Indian interns are less likely to find employment than white interns, even controlling for WEST-B scores or undergraduate GPA.

other omitted variables (i.e., other workforce opportunities) that may be biasing this estimate.

5.2. Internship as screening device: probability of hiring into internship school

Of the 5218 interns hired into the public K-12 system, 806 (15.4%) performed their internship in the building that hired them for their first teaching job. This raises the possibility that schools may use student teaching as a screening process for their own hiring. Our finding that interns who did their student teaching in schools with higher teacher turnover are more likely to be hired lends credence to this notion. We explore this possibility further in Table 4, which reports estimated marginal effects from a logistic regression predicting intern hiring into their internship school (relative to hiring into another school).

The first column of Table 4 reports estimates from a model estimated for all hired interns. In an interesting reversal, non-white interns are more 4.7 percentage points *more* likely to be hired into their internship school than white interns, all else equal. We also experiment with a model that interacts the intern non-white indicator with the percent of non-white students at the internship school, and find that non-white interns who do their student teaching at schools with a high percent of non-white students are particularly likely to be hired by their internship schools.[43] Given that non-white interns are *less* likely to be hired overall, this suggests that non-white interns are particularly unlikely to find a job outside of their internship school. We also find that the probability of employment in the internship school decreases as the stay ratio increases, which matches our hypothesis: interns who do their student teaching at schools with more teacher turnover are more likely to be hired into that school.

Column 2 of Table 4 reports estimates from a model estimated only for interns who did their student teaching at a school that hired at least one new (to the school) teacher the following year (i.e., who had a chance of being hired by their internship school). One intriguing finding from this model is that interns who are endorsed in the same area as their cooperating teacher are more 4.5 percentage points more likely to be hired by their internship school, perhaps reflecting the influence of the cooperating teacher in the *within-school* hiring process.

The last four columns of Table 4 report estimates from models for the four subsets of data we discuss in Section 3: interns with WEST-B scores, undergraduate GPA, the intern's future VAM estimate, or a cooperating teacher VAM estimate (all of whom were hired and did their student teaching at a school that hired at least one new teacher the following year). Interestingly, a ten-point increase in average WEST-B score is correlated with a 2.3 percentage point increase in the probability of being hired by the internship school, and this result is statistically significant. Credential exam scores are modestly correlated with value-added estimates of teacher effectiveness (Goldhaber, 2007), so this result suggests that

schools are more likely to hire their student teachers if they have stronger observable qualifications. The same is not true of GPA, though, and while future intern VAM and cooperating teacher VAM are both positively correlated with the probability of being hired by the internship school, neither result is statistically significant.

6. Discussion and conclusions

In recent years there has been growing attention paid to the role of student teaching in the formulation and progression of an individual's teaching career. Much of this research has investigated the role of student teaching *focusing on individuals who already have become teachers*, thus ignoring the role these internships may play on the decision to become a teacher and their effect on hiring and placement decisions. This paper fills a gap in the literature as it focuses on the relationship between training and workforce entry for a group of *prospective* teachers, arguably among the most important public sector employees since there are over 3.5 million teachers[44], most of whom are trained according using the basic process that we investigate.[45]

We find that the endorsements earned by interns, as well as the characteristics of the schools in which internships take place, are important predictors of whether and when interns are hired into the K-12 system. Interns who receive an endorsement in a STEM field, special education, or ELL are much more likely to be hired into the K-12 system than interns receiving endorsements in other areas. These findings conform to the conventional wisdom that these teachers are in high demand. Moreover, the job market success of these interns suggests that the shortage of STEM and special education teachers may not be the result of inefficiencies in the labor market at time of hire. Rather, shortages in STEM, for example, may be driven by demonstrated differences in the probability that STEM majors pursue teaching degrees (Bacolod, 2007; Goldhaber & Liu, 2003; Ingersoll & Perda, 2010) and by the higher attrition rates of teachers in high-demand areas (Boe, 2006; Fore, Martin, & Bender, 2002; Ingersoll, 2001; Liu, Rosenstein, Swan, & Khalil, 2008; McLeskey, Tyler, & Flippin, 2004).

A related analysis (Goldhaber et al., 2014) finds little evidence that the production of teachers by Washington state teacher training programs has adjusted to the demand for teaching in difficult-to-staff areas. For example, from 1995 to 2010, the ratio of new elementary endorsed teachers produce by TTIs relative to the numbers of job openings was three to two. During this same period, TTI's produced STEM endorsed teachers at only 60% of the rate of STEM job vacancies. This situation has remained relatively stable over the past 30 years, resulting in a surplus of elementary educators and a shortage of STEM

---

[43] Full results are available from the authors on request.

[44] National Center for Education Statistics, Digest of Education Statistics, various tables.

[45] It is worth again emphasizing that we can observe the correlation between intern and internship characteristics and hiring outcomes, but hiring is a two-stage process: a prospective teacher must first decide to pursue a teaching job, and then a school must decide to hire the prospective teacher once he or she has applied. We cannot definitively say whether our results reflect the preferences of prospective teachers or prospective employers.

endorsed teachers. The surplus of elementary endorsements manifests itself in lower probabilities of hire, longer wait times until first job, and an increased likelihood of employment in a private school setting followed by a transition to public school.

Prior research has called into question whether public schools hire more academically talented job applicants (e.g., Ballou, 1996). Our results generally support this conclusion, as we do not observe a strong correlation between licensure exam scores, or grade point average, and the likelihood of being employed in a public teaching position.[46] Interestingly, we do find that interns with higher licensure scores and in internship schools with more teacher turnover are more likely to be hired by their internship schools. This is a previously undocumented route of entry into teaching, and is a relatively common occurrence in our data, accounting for about one-in-six new teacher hires. The frequency of hiring into internship schools suggests that internships serve a dual role: training novice teachers and screening potential new hires. It also raises questions about the placement of interns into internship schools by TTIs; for example, should TTIs place their best interns into schools with higher teacher turnover, knowing that they will be more likely to be hired into these schools?

Some of the non-significant findings are also worth emphasizing.[47] Our research is novel in that we can identify teachers who supervised student internships. Characteristics of these cooperating teachers—such as experience, endorsements, gender, race, educational background and, for a subset of them, value-added—do not appear to be correlated with the probability of an intern's later employment, at least in expected ways. Given the policy interest in improving student teaching (CAEP, 2013; Greenberg et al., 2013) and the perception that student teaching and the quality of the cooperating teacher plays an important role in teacher preparation, these findings are somewhat discouraging as they offer little in the way of direct guidance about how to improve teacher preparation.

The location of a student's internship is also an important determinant in an intern's labor market outcome. Specifically, interns who do their student teaching in suburban schools are more likely to enter the workforce, all else equal. The increased likelihood of being hired from a suburban school may be a result of non-random placement of interns into perceived "healthy" suburban internship schools, a preference for principals to hire current interns, and a

preference for interns to accept positions in suburban schools, all of which would be consistent with the literature showing that teachers tend to be employed near where they did their training (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012). It also suggests that the process by which internships are determined may help explain the distribution of teacher quality across schools, an important topic for future research.

Appendix A. Value-added estimates of teacher effectiveness

Many specifications of our models include an out-of-sample estimate of each cooperating teacher's value-added performance. We refer to the estimates as "out-of-sample" because they are calculated from student test score data from 2005 through 2011, while many internships in our analytic sample fall outside this date range. Other specifications include an "in-sample" estimate of an intern's future (i.e., post-hiring) value-added performance. These estimates are estimated from variants of the following value-added model for all students linked to their classroom teachers in grades 3–8 from 2005 through 2011 in Washington state:[48]

$$Y_{ijst} = \beta_0 + \beta_1 Y_{i(t-1)} + \beta_2 X_{it} + \tau_{js} + \varepsilon_{ijst} \tag{A1}$$

---

[46] This is not terribly surprising given that school systems in Washington (and to our knowledge in other states) do not ask candidates about their scores as part of the teacher application process. Even if licensure scores are predictive of teacher effectiveness, it is possible that they do not strongly correlate with the information collected from teacher applicants at the point of application so would not be correlated with hiring. See Goldhaber (2007) for a focus how licensure scores are used, and on the relationship between licensure scores and teacher effectiveness.

[47] We stress that some statistically insignificant findings, particularly those for models estimated for only a subset of our sample (such as those that include VAM estimates or intern GPA) may not be statistically significant because of low sample sizes. In particular, the hiring into internship school models that include intern or cooperating teacher VAM can be estimated for less than 10% of our overall sample of interns, which explains the high standard errors in these estimates.

---

[48] The proctor of the state assessment was used as the teacher–student link for at least some of the data used for analysis. The 'proctor' variable was not intended to be a link between students and their classroom teachers so this link may not accurately identify those classroom teachers. However, for the 2009–10 school year, we are able to check the accuracy of these proctor matches using the state's new Comprehensive Education Data and Research System (CEDARS) that matches students to teachers through a unique course ID. Our proctor match agrees with the student's teacher in the CEDARS system for about 95 percent of students in math and 94 percent of students in reading. Further, fitting a teacher production function to these data produces similar results to those found elsewhere in the literature (e.g. Clotfelter, Ladd, & Vigdor, 2007).

Y_{ijst} is the state test score for each student i with teacher j in subject s (math or reading) and year t, normalized within grade and year; $Y_{i(t-1)}$ is a vector of the student's scores the previous year in both math and reading, also normalized within grade and year; X_{it} is a vector of student attributes in year t (gender, race, eligibility for free/reduced price lunch, English language learner status, gifted status, special education status, learning disability status, migrant status, and homeless status); and τ_{js} is a fixed effect that captures the contribution of teacher j to student test scores in subject s across all years the teacher is linked to student test score data. We adjust all teacher effect estimates using empirical Bayes (EB) methods.[49]

We use the estimates $\hat{\tau}_{js}$ as a time-invariant measure of a teacher's contribution to student test scores in each subject, math and reading. Since many teachers teach both math and reading, but many secondary teachers only teach math or reading, we use the average of the value-added estimates in math and reading for teachers who teach both subjects. We experiment with variants of model (A1), including models with student and school fixed effects, and find that they do not substantively change our findings.

References

Anderson, L. M., & Stillman, J. A. (2013). Student teaching's contribution to preservice teacher development: A review of research focused on the preparation of teachers for urban and high-needs contexts. *Review of Educational Research, 83*(1), 3–69.

Bacolod, M. (2007). Who teaches and where they choose to teach: College graduates of the 1990s. *Educational Evaluation and Policy Analysis, 29*(3), 155–168.

Ballou, D. (1996). Do public schools hire the best applicants? *Quarterly Journal of Economics, 111*(1), 97–133.

Boe, E. E. (2006). Long-term trends in the national demand, supply, and shortage of special education teachers. *The Journal of Special Education, 40*(3), 138–150.

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., Lankford, H., Loeb, S., Michelli, N., & Wyckoff, J. (2006). Complex by design: Investigating pathways into teaching in New York City schools. *Journal of Teacher Education, 57*(2), 155–166.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Who leaves? Teacher attrition and student achievement (No. w14022)* National Bureau of Economic Research.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 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*(1), 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*(1), 113–132.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2013). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers. *Journal of Labor Economics, 31*(1), 83–117.

Brewer, D. (1996). Career paths and quit decisions: Evidence from teaching. *Journal of Labor Economics, 14*(2), 313–339.

Buxton, A. (2007). *CUREREGR: Stata module to estimate parametric cure regression.*

CAEP: Council for the Accreditation of Educator Preparation (2013). *CAEP commission on standards and performance reporting: Draft recommendations for the CAEP board, Washington D.C..*

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 16*(6), 673–682.

DeAngelis, K. J., & Presley, J. B. (2011). Toward a more nuanced understanding of new teacher attrition. *Education and Urban Society, 43*(5), 598–626.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics, 86*(1), 195–210.

DeYoung, R. (2003). The failure of new entrants in commercial banking markets: A split-population duration analysis. *Review of Financial Economics, 12*(1), 7–33.

Douglas, S., & Hariharan, G. (1994). The hazard of starting smoking: Estimates from a split population duration model. *Journal of Health Economics, 13*(2), 213–230.

Ehrenberg, R., Goldhaber, D., & Brewer, D. (1995). *Do teachers' race, gender, and ethnicity matter? Evidence from NELS88* National Bureau of Economic Research No. w4669.

Engel, M., Jacob, B., & Curran, F. (2014). New evidence on teacher labor supply. *American Educational Research Journal, 51*, 36–72.

Fore, C., Martin, C., & Bender, W. N. (2002). Teacher burnout in special education: The causes and the recommended solutions. *The High School Journal, 86*(1), 36–44.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources, 42*(4), 765–794.

Goldhaber, D. (2013, November). *What do value-added measures of teacher preparation programs tell us?* Carnegie Knowledge Network.

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*(1), 57–87.

Goldhaber, D., Krieg, J., Theobald, R., & Brown, N. (2014). The STEM and special education teacher pipelines: Why don't we see better alignment between supply and demand? *Phi Delta Kappan* (in press, available as CEDR working paper 2014-3).

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.

Goldhaber, D., & Liu, A. (2003). Occupational choices and the academic proficiency of the teacher workforce. In *Developments in School Finance 2001* (Vol. 2, pp. 53–75). .

Goldhaber, D., & Walch, J. (2014). Gains in teacher quality. *Education Next, 14*(1), 38–47.

Greenberg, J., McKee, A., & Walsh, K. (2013, June). *Teacher prep review: A review of the nation's teacher preparation programs*. National Council on Teacher Quality.

Hanushek, E. A., & Pace, R. R. (1995). Who chooses to teach (and why)? *Economics of Education Review, 14*(2), 101–117.

Hess, F. M. (2009). Revitalizing teacher education by revisiting our assumptions about teaching. *Journal of Teacher Education, 60*(5), 450–457.

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534.

Ingersoll, R. M. (2003). *Turnover and shortages among science and mathematics teachers in the United States.*

Ingersoll, R., Merrill, L., & May, H. (2012). Retaining teachers: How teacher preparation matters. *Educational Leadership, 69*(8), 30–34.

Ingersoll, R. M., & Perda, D. (2010). Is the supply of mathematics and science teachers sufficient? *American Educational Research Journal, 47*(3), 563–594.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics, 1*, 85–108.

Jaggia, S. (2011). Identifiability of the misspecified split hazard models. *Applied Economics, 43*(22–24), 3441–3447.

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2014). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance and Policy* (in press, available as University of Missouri Department of Economics Working Paper 12-04).

Krieg, J. M. (2006). Teacher quality and attrition. *Economics of Education Review, 25*(1), 13–27.

Kyyrä, T., & Ollikainen, V. (2008). To search or not to search? The effects of UI benefit extension for the older unemployed. *Journal of Public Economics, 92*(10), 2048–2070.

Liu, E., Rosenstein, J. G., Swan, A. E., & Khalil, D. (2008). When districts encounter teacher shortages: The challenges of recruiting and retaining mathematics teachers in urban districts. *Leadership and Policy in Schools, 7*(3), 296–323.

---

[49] The standard empirical Bayes method shrinks estimates back to the grand mean of the population. Note, however, that standard empirical Bayes adjustment does not properly account for the uncertainty in the grand mean, suggesting the estimates are shrunk too much (McCaffrey, Sass, Lockwood, & Mihaly, 2009). We use the standard approach that has been commonly estimated in the literature (an appendix on empirical Bayes shrinkage is available from the authors on request).

McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy, 4*(4), 572–606.

Maggiolini, P., & Mistrulli, P. E. (2005). A survival analysis of de novo co-operative credit banks. *Empirical Economics, 30*(2), 359–378.

Maier, A., & Youngs, P. (2009). Teacher preparation programs and teacher labor markets: How social capital may help explain teachers' career choices. *Journal of Teacher Education, 60*(4), 393–407.

McLeskey, J., Tyler, N. C., & Flippin, S. S. (2004). The supply of and demand for special education teachers: A review of research regarding the chronic shortage of special education teachers. *The Journal of Special Education, 38*(1), 5–21.

Mihaly, K., McCaffrey, D., Sass, T., & Lockwood, J. R. (2013). Where you come from or where you go? Distinguishing between school quality and the effectiveness of teacher preparation program graduates. *Education Finance and Policy, 8*(4), 459–493.

Papay, J. P., West, M. R., Fullerton, J. B., & Kane, T. J. (2012). Does practice-based teacher preparation increase student achievement? Early evidence from the Boston Teacher Residency. *Educational Evaluation and Policy Analysis, 34*(4), 413–434.

Podgursky, M., Monroe, R., & Watson, D. (2004). The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review, 23*(5), 507–518.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from. Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis, 34*(2), 127–145.

Ronfeldt, M. (2012). Where should student teachers learn to teach? Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis, 34*(1), 3–26.

Ronfeldt, M., Schwartz, N., & Jacob, B. (2014). Does pre-service preparation matter? Examining an old question in new ways. *Teachers College Record, 116*(10), 1–46.

Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review, 26*(2), 145–159.

Schmidt, P., & Witte, A. D. (1989). Predicting criminal recidivism using 'split population' survival time models. *Journal of Econometrics, 40*(1), 141–159.

Stinebrickner, T. R. (2001). Compensation policies and teacher decisions. *International Economic Review, 42*(3), 751–780.

Swaim, P., & Podgursky, M. (1994). Female labor supply following displacement: A split-population model of labor force participation and job search. *Journal of Labor Economics,* 640–656.

White, B. R., DeAngelis, K. J., & Lichtenberger, E. J. (2013). *The student has become the teacher: Tracking the racial diversity and academic composition of the teacher supply pipeline.* Edwardsville, IL: Illinois Education Research Council at Southern Illinois University Edwardsville.