



NATIONAL
CENTER *for* ANALYSIS of LONGITUDINAL DATA in EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington



How Did It Get This Way? Disentangling the Sources of Teacher Quality Gaps Across Two States

Dan Goldhaber
Vanessa Quince
Roddy Theobald

How Did It Get This Way? Disentangling the Sources of Teacher Quality Gaps Across Two States

Dan Goldhaber

American Institutes for Research/ CALDER

University of Washington

Vanessa Quince

University of Washington

Roddy Theobald

American Institutes for Research/ CALDER

Contents

Contents.....	i
Acknowledgment.....	ii
Abstract.....	iii
1. Introduction.....	1
2. Framework and Literature Review.....	4
3. Data.....	8
4. Notation and Summary Statistics	13
5. Simulation	17
6. Results.....	20
7. Conclusions.....	25
References.....	28
Tables & Figures.....	32

Acknowledgments

This work is supported by the William T. Grant Foundation (grant #184925) and by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see www.caldercenter.org/about-calder. All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders or the institutions to which the authors are affiliated.

The research presented here is based in part on confidential data from the North Carolina Education Research Center (NCERDC) at Duke University, directed by Kara Bonneau. We wish to acknowledge the North Carolina Department of Public Instruction for its role in collecting this information and making it available, as well as the Washington State Office of the Superintendent of Public Instruction (OSPI) for collecting the student-level data from Washington State and sharing it through data-sharing agreement No. 2015DE-028. We wish to thank Joe Walch and Malcolm Wolff for research assistance, as well as Chad Aldeman, James Cowan, Aliza Husain, and Jason Grissom for comments that improved the manuscript.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any opinions, findings, and conclusions expressed in these papers are those of the authors and do not necessarily reflect the views of our funder or the data providers.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street NW, Washington, DC 20007
202-403-5796 • www.caldercenter.org

How Did It Get This Way? Disentangling the Sources of Teacher Quality Gaps Across Two States

Dan Goldhaber, Vanessa Quince, Roddy Theobald
CALDER Working Paper No. 209-1118-1
November 2018

Abstract

We use longitudinal data from North Carolina and Washington to study the extent to which four processes—teacher attrition from each state workforce, teacher mobility within districts, teacher mobility across districts, and teacher hiring—contribute to “teacher quality gaps” (TQGs) between advantaged and disadvantaged schools. We first replicate prior findings documenting inequities in each of these processes using different measures of student disadvantage (race and poverty) and teacher quality (experience, licensure test scores, and value added) and then develop and implement a simulation to assess the extent to which each process contributes to observed TQGs in each state. We find that all four processes contribute to TQGs but also document considerable heterogeneity in the extent to which each process contributes to the different TQG measures. For example, patterns in teacher attrition and mobility contribute more to TQGs measured by teacher experience, while patterns in teacher hiring explain the majority of TQGs measured by teacher licensure test scores and value added.

1. Introduction

There is mounting descriptive evidence of teacher quality gaps (TQGs) between advantaged and disadvantaged students in U.S. public schools. These TQGs are evident whether teacher quality is measured by degrees, experience, or advanced credentials (e.g., Clotfelter, Ladd, & Vigdor, 2005; Kalogrides & Loeb, 2013; Lankford, Loeb, & Wyckoff, 2002) or by “value-added” measures of teacher effectiveness (e.g., Goldhaber, Lavery, & Theobald, 2015; Isenberg et al., 2016; Sass, Hannaway, Xu, Figlio, & Feng, 2012). Recent evidence from longitudinal data on public schools from North Carolina and Washington (Goldhaber, Quince, & Theobald, 2018) demonstrates that TQGs have existed in every available year of data in each state and for each observable measure of student disadvantage (i.e., race/ethnicity and poverty level) and teacher quality (i.e., experience, licensure test scores, and value added).

In this paper, we assess the extent to which four different processes—the attrition of teachers from each state workforce, the movement of teachers between schools within a district, the movement of teachers between districts, and the hiring of teachers into open teaching positions—contribute to TQGs between advantaged and disadvantaged schools. Consistent with prior literature, we document inequitable patterns in each state (North Carolina and Washington) in the relationship between each measure of teacher quality (experience, licensure test scores, and value added) and the four process that can generate TQGs.

Our primary contribution, however, is to disentangle the extent to which each process contributes to overall TQGs. Understanding the sources of TQGs is fundamental to closing them, a policy goal that has been elevated by the recent directive to states from the U.S. Department of Education is to develop plans to reduce inequity in the distribution of teacher quality across public schools (Rich, 2014). For example, if most of the inequity between advantaged and

disadvantaged schools is the result of differential teacher attrition, states and districts might focus on policies to *keep* high-quality teachers in disadvantaged schools. If, on the other hand, discrepancies in teacher hiring explain most of the observed inequities, states and districts might instead prioritize policies to *attract* high-quality teachers to disadvantaged schools.¹

The heart of our analysis is the implementation of a stochastic model of the proportion of low-quality teachers in advantaged and disadvantaged schools that is a function of all four processes described above.² We use this model to simulate a number of different scenarios—that is, to simulate scenarios in which each of the processes above is the *only* source of inequity in the attrition, movement, and hiring of teachers—that allow us to examine the extent to which each process independently contributes to TQGs between advantaged and disadvantaged students.

The stochastic model we use is relatively simple in that it classifies each school as either “advantaged” or “disadvantaged” based on the percentage of disadvantaged students in the school; in addition, schools do not change status over time. We also make the simplifying assumptions that all teachers in the simulation are either “high quality” or “low quality.” Finally and most importantly, the model we use ignores the potential that one process might affect another. For example, we do not consider whether the attrition of high- or low-quality teachers from a school is impacted by the type of teachers that were previously hired into the school. This is an important limitation in that teachers respond to the quality of their peers in making mobility decisions (Feng & Sass, 2017), and there is evidence that there are teacher quality spillovers and

¹ Of course, it might be most cost-effective to address TQGs by creating inequity in one process to offset inequity in another process. For instance, disadvantaged schools might reduce the attrition of high-quality teachers below the level of attrition of high-quality teachers in advantaged schools to offset gaps in the ability of the two types of school to hire high-quality teachers. Assessing the costs of addressing TQGs is outside the scope of this paper, but we also believe that the politics of implementing solutions to TQGs are likely connected to the sources of the gaps, suggesting the need to better understand these.

² We use the term “stochastic model” to reference the fact that our simulation is an example of a stochastic process (or, more precisely, a Markov process) in which outcomes in year t are solely a function of outcomes in year $t-1$ and transition probabilities between these outcomes.

in particular that the value added of teachers is affected by the quality of their teacher peers (Jackson & Bruegmann, 2009).

Yet despite these limitations, we believe this exercise is useful because it provides information to policymakers about which processes appear to be contributing most to TQGs and therefore where they might intervene to close these TQGs. As a specific example, suppose a policymaker is looking to close TQGs by offering recruitment or retention bonuses to high-quality teachers. The success of these interventions depends both on the extent to which the intervention impacts the process it is intended to impact (hiring and attrition, respectively) *and* on the extent to which making these processes more equitable might close TQGs. Much of the existing literature focuses exclusively on the first mechanism, while this paper is intended to address the second.

For nearly every combination of state, measure of teacher quality, and measure of student disadvantage, we find that *all four processes* contribute to TQGs.³ But we also find considerable heterogeneity in the extent to which each process contributes to TQGs by state and measure of teacher quality. For example, teacher attrition and mobility contribute more to TQGs in North Carolina than in Washington, while teacher hiring contributes more to TQGs in Washington than in North Carolina. And in both states, teacher attrition and mobility contribute more to TQGs measured by teacher experience, while teacher hiring explains the lion's share of TQGs measured by teacher licensure test scores and value added.

The fact that teacher attrition and mobility contribute substantially to TQGs as measured by teacher experience suggests that policymakers seeking to improve equity in access to experienced teachers should focus on *retention* policies, as is often recommended in policy

³ One notable exception is that we generally do not find that patterns in teacher attrition contribute to TQGs in terms of teacher value added.

circles (e.g., Carver-Thomas & Darling-Hammond, 2017). The importance of teacher hiring in TQGs with respect to licensure tests and value added, on the other hand, suggests that *recruitment and hiring* policies may be an underexplored lever for closing these TQGs. This recommendation is tempered somewhat by an extension in which we show that gaps in hiring by teacher licensure test scores have been extremely persistent over time in both states, which suggests that these inequities may be less malleable than other inequities we study. That said, we need to learn much more about *why* teacher hiring appears to be so inequitable to fully contextualize these findings.

The paper proceeds as follows. In Section 2, we discuss our framework for investigating each of the labor market processes discussed above and the prior literature related to each process. We describe the data in Section 3, outline our notation and present summary statistics in Section 4, develop the simulations in Section 5, and discuss the results of the simulations in Section 6. We conclude with implications for policy and directions for future research in Section 7.

2. Framework and Literature Review

A growing literature documents the existence and magnitude of TQGs in districts and states across the country (see Goldhaber et al., 2018, for a review), but perhaps equally important is how these inequities formed in the first place. A recent report from the U.S. Department of Education (Reform Support Network, 2015) identifies a number of processes within the teacher pipeline—the attrition of teachers from different types of schools, the movement of teachers between schools and districts, and the hiring of new teachers into their first jobs—that could

contribute to TQGs across U.S. public schools. The literature on the teacher labor market has documented trends in each of these individual processes that may contribute to TQGs.

In this section, we discuss the prior research on four mutually exclusive processes that could potentially contribute to TQGs.

Teacher attrition. A number of multistate and national studies demonstrate that teachers in disadvantaged schools are more likely to leave the workforce (e.g., Grissom, 2011; Ingersoll & May, 2012; Kaiser, 2011; Keigher, 2010; Marvel et al., 2007; Shen, 1997). For instance, in a study of 26 school districts, Isenberg et al. (2016) find that 10% of teachers leave high-poverty schools in a typical year compared to 7% of teachers who leave low-poverty schools. Studies that focus on specific schools and districts tend to come to similar conclusions. For example, Barnes (2007) finds a 6% difference in turnover between high-poverty and low-poverty schools in Chicago and Illinois public schools, while Cook (2011) finds that a 25% increase in the proportion of students eligible for free or reduced-price meals is associated with a .20 percentage point increase that a teacher will leave the teaching profession in North Carolina. These observed patterns are consistent with findings from various individual states like Florida (Feng & Sass, 2017; Ingle, 2009), Georgia (Scafidi, Sjoquist, & Stinebrickner, 2007), New York (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2008), Texas (Hanushek, Kain, & Rivken, 2004), Wisconsin (Imazeki, 2005), and our focal states of North Carolina (Clotfelter, Glennie, Ladd, & Vigdor, 2008; Goldhaber, Gross, & Player, 2011) and Washington (Goldhaber, Quince, & Theobald, 2016; Gritz & Theobald, 1996; Krieg, 2006).

The above literature shows that teacher attrition is generally lower in advantaged schools, but evidence of differential attrition by teacher effectiveness (measured by value added) suggests that the role attrition plays in some types of TQGs is more complicated. Specifically, Krieg

(2006), Goldhaber et al. (2011), Isenberg et al. (2016), and Feng and Sass (2017) all find that ineffective teachers are disproportionately likely to leave the workforce, and Goldhaber et al. (2011) and Isenberg et al. (2016) report that this relationship is stronger in disadvantaged schools than advantaged schools. Thus, whether attrition increases or decreases value-added TQGs depends on the effectiveness of those teachers who are leaving different types of schools.

Within-district mobility. The empirical evidence on within-district teacher mobility suggests that within-district transfers may contribute to TQGs in two ways. First, teachers who teach in schools with higher proportions of disadvantaged students are more likely to transfer to another school in the same district than those who teach in schools with lower proportions of disadvantaged students (e.g., Cook 2011; Goldhaber et al., 2011; Hanushek et al., 2004; Isenberg et al., 2016; Sass et al., 2012; Scafidi et al., 2007). For example, Isenberg et al. (2016) find that on average 11% of teachers in high-poverty schools transfer to another school in the district compared to 5% in low-poverty schools. This can contribute to TQGs because each year disadvantaged schools must hire more new teachers, who on average are less experienced and less effective than the average teacher (e.g., Isenberg et al., 2016).

Second, within-district mobility can also contribute to teacher quality gaps if high-quality teachers are disproportionately more likely to leave disadvantaged schools (or, alternatively, that low-quality teachers are disproportionately likely to stay in disadvantaged schools). There is some prior evidence of this (e.g., Boyd et al. 2008; Goldhaber et al., 2011; Hanushek and Rivkin 2010; Isenberg et al., 2016); for example, Isenberg et al. (2016) find that effective teachers (as measured by value added) are more likely than ineffective teachers to leave high-poverty schools for low-poverty schools in the same district.

Cross-district mobility. Cross-district mobility can contribute to TQGs in the same ways as within-district mobility. Teacher mobility across districts has received less empirical attention, likely because analysis requires access to statewide databases, but prior work on cross-district mobility also suggests teachers are more likely to leave disadvantaged schools for a school in a different district (Goldhaber et al., 2011; Hanushek et al., 2004; Scafidi et al., 2007). Moreover, the one study we are aware of that investigates differences in cross-district mobility by teacher effectiveness and school disadvantage (Goldhaber et al., 2011) finds that effective teachers (as measured by value added) are disproportionately likely to leave schools with a high percentage of minority students for a school in a different district.

Teacher hiring. The patterns described above imply that, on average, disadvantaged schools need to hire more teachers in a given year than advantaged schools. In our framework, we attribute any inequities that arise from these *overall* higher levels of teacher hiring to the processes discussed above. On the other hand, patterns in teacher hiring can further contribute to TQGs if, *conditional on a position being open*, disadvantaged schools and districts are less likely to hire an effective teacher. Empirical evidence in this area is mixed; some studies find that there are no discernable differences between the value added of teachers hired to high-poverty schools and those hired to low-poverty schools (Isenberg et al., 2016; Sass et al., 2012), while Xu, Ozek, & Hansen, (2015) find that teachers in their first 2 years at high-poverty schools were 0.02 standard deviations lower in terms of value added than teachers at low-poverty schools, and Loeb, Kalogrides, & Béteille, (2012) find that effective schools are more likely to hire teachers with higher prior estimates of value added.⁴ We are not aware of empirical evidence about the

⁴ Importantly, these papers do not focus specifically on teacher hiring but simply compare the effectiveness of novice teachers in different school settings.

conditional probability of hiring a novice teacher or a teacher with a low licensure test score in different types of schools, which is one unique contribution of this paper.

3. Data

3.1. Data Overview

For our analysis, we combine administrative data from North Carolina, provided by the North Carolina Education Research Data Center (NCERDC), and data from Washington, provided by the Washington State Office of Superintendent of Public Instruction (OSPI), along with data from the National Center for Education Statistics (NCES). The measures of teacher quality come from the state administrative databases, described more extensively below. As in Goldhaber et al. (2018), we focus on the distribution of “low-quality” teachers: novice teachers (teachers with 5 or fewer years of experience), teachers in the lowest quartile of the distribution of licensure test scores, and teachers in the lowest quartile of the distribution of value added. Existing literature suggests that exposure to low-quality teachers is important for all three measures of teacher quality. This relationship has been documented most strongly for teacher experience (e.g., Rice, 2013; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), but, more recently, evidence suggests that teachers in the bottom quartile of licensure test scores drive the relationship between teacher licensure test scores and student achievement as well (Goldhaber, Gratz, & Theobald, 2017).

The data about school disadvantage are derived from the Public School Universe survey maintained by the NCES, which includes school-level data about the percentage of students by race and ethnicity as well as the percentage of economically disadvantaged (ED) students.⁵ Our

⁵ In response to a directive from the North Carolina Education Research and Data Center, we use the term “economically disadvantaged students” to refer to students who qualify for free or reduced-price meals.

primary measures of school disadvantage are indicators for whether a school falls in the top quartile of the distribution of the percentage of underrepresented minority (URM) students—American Indian, Black, or Hispanic—or whether a school falls in the top quartile of the distribution of the percentage of ED students. While there is greater availability in data concerning the race and ethnicity of students, for the sake of consistency we begin our analysis with the 1999 school year (since this is the first year with non-missing ED data).⁶

3.2. North Carolina Data

The data set in North Carolina uses teacher-level data from the NCERDC spanning 15 school years (1998–99 through 2012–13). The data set includes information on teacher positions, salary, and teaching experience; we limit this data set to teachers with a full-time teaching appointment in a single school within a given year. The data set also includes information on teachers' licensure test scores in reading, writing, and math on the state's Praxis teacher licensure test.⁷ We begin our Praxis analysis with the 1999–2000 school year since it is the first year where at least 1% of teachers in the state have a score. We consider the average of each teacher's score on the math, reading, and writing portions of the exam for the *first* time each teacher took the test.⁸

⁶ These school-level measures are highly correlated with school-level measures of academic performance. Specifically, the correlation between school percent URM and school average math performance is -0.62 in North Carolina and -0.48 in Washington; the correlation between school percent URM and school average reading performance is -0.70 in North Carolina and -0.61 in Washington; the correlation between school percent EDS and school average math performance is -0.77 in North Carolina and -0.66 in Washington; and the correlation between school percent EDS and school average reading performance is -0.83 in North Carolina and -0.74 in Washington.

⁷ In North Carolina, teachers must also pass subject assessment tests, but those exams were not included in the analysis.

⁸ Teachers may take licensure tests multiple times to get a passing score on all three tests, so we use the test scores from the first time each teacher took the Praxis (and follow a similar procedure with the WEST-B in Washington). This ensures that teachers taking the test for the fifth time, for example, are not judged as “comparable” to teachers who passed all three tests on the first attempt.

For math and reading teachers in grades 4–6, we include an estimate of a teacher’s effectiveness calculated from the variants of the following value-added model (VAM) estimated separately for both math and reading:

$$Y_{ijst} = \beta_0 + \beta_1 Y_{i(t-1)} + \beta_2 S_{it} + \beta_3 C_{jt} + \tau_{js} + \varepsilon_{ijst} \quad (1)$$

In equation 1, 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; S_{it} is a vector of student attributes in year t (gender, race, ED, English language learner status, gifted status, special education status, learning disability status); C_{jt} is a vector of these student attributes aggregated to the classroom level; and τ_{js} is the VAM estimate that captures the contribution of teacher j to student test scores in subject s up to and including year t .⁹

We estimate the above value-added specification both with and without the vector of classroom-level student controls C_{jt} . There is debate in the literature on TQGs about which specification is preferable for estimating TQGs in terms of value added (e.g., Goldhaber et al., 2016; Isenberg et al., 2016), so we report all results in this paper for both specifications. The estimates without classroom controls facilitate direct comparisons with our prior published work (e.g., Goldhaber et al., 2015, 2018), while the estimates with classroom controls facilitate comparisons with other work (e.g., Isenberg et al., 2016). Importantly, this paper focuses on elementary grades in estimating value added, and prior work (Goldhaber et al., 2016) suggests that these two specifications lead to similar conclusions about TQGs at these grade levels.

⁹ Because of computing limitations, we consider only up to 7 years of prior data in estimating these VAMs in North Carolina.

Because we are primarily interested in differences in exposure for teachers at the tail of the value-added distribution and studies demonstrate that teachers matched to smaller numbers of students are more likely to be in the tails of the distribution (e.g., Aaronson, Barrow, & Sander, 2007), we improve the precision of our estimates for each year t by focusing on pooled value-added estimates that consider all available years of data up to and including year t for each teacher.¹⁰ Additionally, we adjust all teacher effect estimates using empirical Bayes (EB) methods to shrink the estimates back to the grand mean of the value-added distribution in proportional to the standard error of each estimate. EB shrinkage does not account for the uncertainty in the grand mean, suggesting that estimates may shrink too much under this procedure (McCaffrey, Sass, Lockwood, & Mihaly, 2009); this approach, however, ensures that estimates in the tail of the distribution are not disproportionately estimated with large standard errors. We use the average math and reading value-added estimates for teachers who teach both subjects.¹¹

The merged North Carolina NCES data include 1,172,723 teacher-year observations (204,949 unique teachers) spanning 15 school years. We observe Praxis scores for 165,402 of these teacher-year observations and value-added estimates for 207,747 teacher-year observations.¹²

3.3. *Washington Data*

For Washington, we use the state's S-275 database, which contains information from OSPI's personnel-reporting process and includes school assignment of all certified employees in

¹⁰ This increased precision comes at the cost of ignoring true changes in teacher quality over time.

¹¹ We also experiment with additional specifications of the model in equation 1, including a model that has indicators for teacher experience level (so comparisons are made of students assigned to teachers with the same teaching experience), a model that controls for 2 years of prior test scores, and a model that corrects for measurement error in the prior test scores.

¹² Specifically, we observe Praxis scores for teachers who were credentialed since the Praxis became a requirement and value added for teachers who teach math or reading in grades 4–6.

the state in addition to a measure of teaching experience in the state.¹³ S-275 data are available from the 1983–84 school year through the 2016–17 school year, although our analysis in Washington spans 13 years (2001–02 through 2013–14) because 2001–02 is the first year of non-missing ED data in the NCES data set for Washington and 2013–14 was the last year of data considered in our prior analysis of TQGs (Goldhaber et al., 2018). As in North Carolina, we consider teachers with a full-time teaching appointment in a single school within a given year.

We link the S-275 data set to the same teacher quality measures described above for North Carolina. First, we consider teacher experience and specifically whether a teacher has fewer than 5 (or, in extensions, fewer than 2) years of experience. Second, we include the teacher’s test score on the Washington Educator Skills Test – Basic (WEST-B). WEST-B is the standardized test that all teachers must pass before entering a teacher education program. As in North Carolina, we consider average scores across reading, writing, and math for the first time each teacher took the test. Since the WEST-B was not required until 2002, our analysis considers teacher WEST-B scores in the 2005–06 school year because it is the first year where at least 1% of teachers have a WEST-B score. Lastly, we use the same specification described in equation 1 to estimate teacher value added for teachers in grades 4–6 going back to the 2006–07 school year (the first year in which current and prior test scores are available in these grades).

After merging with the NCES data, the final longitudinal data set in Washington includes 564,296 teacher-year observations (86,241 unique teachers) spanning 13 school years. We observe WEST-B scores for 69,996 of these teacher-year observations and value-added estimates from 41,416 teacher-year observations.

¹³ The S-275 contains the experience that teachers are credited with for pay purposes, which may not include out-of-state teaching, teaching in a private school, or substitute teaching.

4. Notation and Summary Statistics

4.1. Assumptions and Notation

We begin this section by outlining the assumptions we make for the simulation and the notation we use throughout the rest of the paper. As described in Section 3, we consider dichotomous measures of school disadvantage; that is, schools are either advantaged (ADV) or disadvantaged (DIS) and teachers are either low quality (LQ) or high quality (HQ).¹⁴ We also simplify our data by averaging the percentage of disadvantaged students (%URM and %EDS) in a school across all years of available data; this makes a school's disadvantaged classification time invariant. Importantly, however, it is likely to have little impact on our analysis given that the year-to-year pairwise correlation in the percentage of disadvantaged students is over 0.85 in each state and for each measure of student disadvantage.

Next, we define the measures of interest for each school type $S \in \{ADV, DIS\}$ and teacher quality category $Q \in \{LQ, HQ\}$. First, define $p_{Q,t}^S$ as the proportion of teachers of quality Q in a school of type S in year t . We use these probabilities to define the teacher quality gap in year t , TQG_t , as the difference in the proportion of low-quality teachers between disadvantaged and advantaged schools:

$$TQG_t = p_{LQ,t}^{DIS} - p_{LQ,t}^{ADV} \quad (2)$$

Next, we define probabilities associated with six mutually exclusive outcomes for teachers after a given year t :

¹⁴ We plan to consider continuous measures of school disadvantage and teacher quality in future work.

$p_{Q,t}^{STAY,S}$ = The proportion of teachers of quality Q in a school of type S who stay in the same school after year t .

$p_{Q,t}^{WD,S \rightarrow S}$ = The proportion of teachers of quality Q in a school of type S who move within a district to another school of type S after year t .

$p_{Q,t}^{WD,S \rightarrow S'}$ = The proportion of teachers of quality Q in a school of type S who move within a district to another school of type S' after year t .

$p_{Q,t}^{BD,S \rightarrow S}$ = The proportion of teachers of quality Q in a school of type S who move between districts to another school of type S after year t .

$p_{Q,t}^{BD,S \rightarrow S'}$ = The proportion of teachers of quality Q in a school of type S who move between districts to another school of type S' after year t .

$p_{Q,t}^{ATT,S}$ = The proportion of teachers of quality Q in a school of type S who leave the workforce after year t .

Note that the six above outcomes encompass all possible outcomes for a teacher after year t , so

$$p_{Q,t}^{STAY,S} + p_{Q,t}^{WD,S \rightarrow S} + p_{Q,t}^{WD,S \rightarrow S'} + p_{Q,t}^{BD,S \rightarrow S} + p_{Q,t}^{BD,S \rightarrow S'} + p_{Q,t}^{ATT,S} = 1.$$

Finally, conditional on an open position after teacher attrition and mobility after year t , we define hiring proportions in year $t+1$:

$p_{Q,t+1}^{HIRE,S}$ = The proportion of teachers hired into schools of type S in year $t+1$ who are of quality Q .

We note that, unlike the other proportions outlined above, the interpretation of the hiring proportions varies considerably depending on the measure of teacher quality we consider.

Specifically, while the probability of hiring a novice teacher into an open position is

straightforward, the probability of hiring a low-value-added teacher is conditional on hiring a teacher who has a prior value-added estimate.

In Section 5, we derive the TQG in year $t+1$ as a function of the proportions described above, which forms the basis of our stochastic simulation. In the next section, we calculate and present these probabilities from the data sets described in Section 4.

4.2. *Summary Statistics*

We present summary statistics of the key measures considered in this study in **Figures 1–8**. The panels in Figures 1–8 correspond to the proportions described in Section 4.1; these proportions are calculated across all years of available data in each state (e.g., we consider the mobility of teachers in Washington between 2002 and 2012 and the hiring of teachers between 2003 and 2013 and average these proportions across all years of data).¹⁵ Note that the mobility of teachers between schools of the same level of disadvantage S has no impact on teacher quality gaps,¹⁶ so we focus on just four sets of proportions from Section 4.1: the attrition proportions $p_{Q,t}^{ATT,S}$ (Panel A),¹⁷ the within-district mobility proportions $p_{Q,t}^{WD,S \rightarrow S'}$ (Panel B), the between-district mobility proportions $p_{Q,t}^{BD,S \rightarrow S'}$ (Panel C), and the hiring proportions $p_{Q,t+1}^{HIRE,S}$ (Panel D).

Each of the figures considers a different combination of student disadvantage measure and teacher quality measure. Many of the summary statistics in these figures reinforce the conclusions from the existing literature described in Section 2. For example, Panel A in every

¹⁵ We plan to consider variation in these proportions across different years in future work.

¹⁶ Policymakers may care about this type of mobility for other reasons; for example, empirical evidence suggests that the “churn” of teachers between schools is detrimental to student achievement independent of the effectiveness of entering and departing teachers (Ronfeldt et al., 2013).

¹⁷ Because of our definition of teachers (i.e., individuals in a full-time teaching position in a given school and year), teachers can be counted as “leaving the workforce” if they are no longer full-time instructors but are still within the public school system, move into another role within the school such as an administrator or school principal, or take a temporary leave the following year.

figure illustrates that teacher attrition is more prevalent in disadvantaged schools in each state, and Panels B and C illustrate that teachers are more likely (than we would expect by random chance)¹⁸ to move from disadvantaged schools to advantaged schools than vice versa. There is also some evidence of *differential* mobility and attrition that reinforces prior findings; for example, Panel A of Figure 1 illustrates that non-novice teachers are disproportionately likely to leave advantaged schools, while Panel A of Table 5 demonstrates that bottom quartile value-added teachers are both more likely to leave the workforce than more effective teachers and are disproportionately likely to leave disadvantaged schools.

The above findings largely reflect the findings described in the literature discussed in Section 2. More novel are the findings for teacher hiring and for teacher licensure tests as a measure of quality. In particular, the hiring proportions (Panel D in Figures 1–8) illustrate that, conditional on an open position, disadvantaged schools are more likely to hire a low-quality teacher than advantaged schools, regardless of how we define teacher quality and student disadvantage. We are not aware of prior work that has illustrated this specific source of inequity for teacher experience and licensure test scores, while our findings for value added echo similar findings from Xu et al. (2015). Moreover, we are not aware of prior work that considers variation in teacher attrition, mobility, and hiring by teacher licensure test scores (Figures 3 and 4). One striking conclusion from these figures is that teacher hiring is particularly inequitable when it comes to the distribution of teacher licensure test scores; for example, as shown in Panel D of Figure 4, high-URM schools in North Carolina are about 50% more likely to hire a teacher with a bottom quartile licensure test score than lower-URM schools in the state.

¹⁸ Because there are about 3 times as many teachers in advantaged schools as disadvantaged schools according to our definition, we would expect by random chance that teachers would be 3 times more likely to move from disadvantaged to advantaged schools than vice versa. Panels B and C in Figures 1–6 illustrate that these proportions are typically 4–6 times greater.

5. Simulation

5.1. Stochastic Model

We now use the proportions defined in Section 4.1 and summarized in Section 4.2 to derive a stochastic model that we will use in the simulations, described in Section 5.2, that are intended to disentangle the extent to which each of these processes contributes to TQGs in each state. First, we define $N_{Q,t}^S$ as the number of teachers of quality Q in a school of type S in year t and use this to define the number of teachers associated with each probability $p_{Q,t}^{*,S}$ defined in Section 4.1.

$$n_{Q,t}^{*,S} = p_{Q,t}^{*,S} * N_{Q,t}^S \quad (3)$$

We now use these definitions to derive the number of low-quality teachers ($Q = LQ$) in advantaged and disadvantaged schools in year $t+1$. To make the derivation clear, we derive the number of low-quality teachers in advantaged schools in year $t+1$ ($N_{LQ,t+1}^{ADV}$) in two steps: (a) the number of *returning* teachers of quality $Q \in \{LQ, HQ\}$ in advantaged schools ($S = ADV$) in year $t+1$, $N_{Q,t+1}^{ADV,RETURN}$, and (b) the number of *new* low-quality teachers in advantaged schools in year $t+1$, $N_{LQ,t+1}^{ADV,NEW}$. First, the number of returning teachers in advantaged schools is the number of teachers who stay in the same advantaged school in year $t+1$ plus the number of teachers who transfer to an advantaged school in year $t+1$:

$$N_{LQ,t+1}^{ADV,RETURN} = n_{LQ,t}^{STAY,ADV} + n_{LQ,t}^{WD,ADV \rightarrow ADV} + n_{LQ,t}^{WD,DIS \rightarrow ADV} + n_{LQ,t}^{BD,ADV \rightarrow ADV} + n_{LQ,t}^{BD,DIS \rightarrow ADV} \quad (4)$$

$$N_{HQ,t+1}^{ADV,RETURN} = n_{HQ,t}^{STAY,ADV} + n_{HQ,t}^{WD,ADV \rightarrow ADV} + n_{HQ,t}^{WD,DIS \rightarrow ADV} + n_{HQ,t}^{BD,ADV \rightarrow ADV} + n_{HQ,t}^{BD,DIS \rightarrow ADV} \quad (5)$$

Second, the number of new low-quality teachers in advantaged schools in year $t+1$ can be calculated as the number of open positions in advantaged schools in year $t+1$ —that is, the

number of teachers in advantaged schools in year t who do *not* return to an advantaged school in year $t+1$ —times the probability that an advantaged school hires a low-quality teacher to fill an open position:

$$N_{LQ,t+1}^{ADV,NEW} = (N_{LQ,t}^{ADV} + N_{HQ,t}^{ADV} - N_{LQ,t+1}^{ADV,RETURN} - N_{HQ,t+1}^{ADV,RETURN}) * p_{LQ,t+1}^{HIRE,ADV} \quad (6)$$

The number of low-quality teachers in advantaged schools in year $t+1$, then, can be calculated as

$$N_{LQ,t+1}^{ADV} = N_{LQ,t+1}^{ADV,RETURN} + N_{LQ,t+1}^{ADV,NEW} \quad (7)$$

An analogous calculation gives the number of low-quality teachers in disadvantaged schools in year $t+1$, $N_{LQ,t+1}^{DIS}$. Thus equations 3–7 provide a closed-form set of equations that take each of the probabilities defined in Section 4.1 and the number of high-quality and low-quality teachers in advantaged and disadvantaged schools in year t and calculate the number of high-quality and low-quality teachers in advantaged and disadvantaged schools in year $t+1$, which can in turn be used to calculate the TQG in year $t+1$ (see equation 2 in Section 4.1). We call this a stochastic model because, like a stochastic process (or more accurately, a Markov process), it is not responsive to time-variant factors that may influence teacher mobility or attrition (e.g., labor market factors, the influence of peer decisions, etc.) but rather takes a set of time-invariant probabilities (discussed in Section 4) and calculates how TQGs change as a function of these probabilities.

5.2. *Simulation*

The stochastic model described in Section 5.1 allows us to connect patterns in teacher attrition, mobility, and hiring directly to TQGs. We use this model to perform a number of simulations that seek to isolate the contribution of each process described earlier (teacher attrition, within-district teacher mobility, cross-district teacher mobility, and teacher hiring) to

TQGs. Because it is difficult to isolate the contribution of each of these processes from other factors that have influenced *existing* TQGs, each of our simulations begins by “hitting the reset button” and assuming that the distribution of teacher quality is perfectly equitable between advantaged and disadvantaged schools in the base year of the simulation (i.e., $TQG_{t=0} = 0$).

For each combination of state (North Carolina or Washington), measure of teacher quality (experience, licensure test scores, or value added), and measure of school disadvantage (%URM or %EDS)—12 combinations in all—we then perform four different simulations. In the first simulation, we assume that teacher attrition is the *only* source of inequity in the teacher labor market; we do this by setting rates of teacher attrition to their observed values calculated from the data (see Section 4.2 and Figures 1–8) but then setting rates of within-district mobility, cross-district mobility, and hiring to be the same in advantaged and disadvantaged schools. In the second simulation, we assume that within-district mobility is the only source of inequity. Cross-district mobility is the only source of inequity in the third simulation, and teacher hiring is the only source of inequity in the fourth simulation. We include additional details about each of these simulations in the appendix.

The intuition behind these simulations is simple: If we’re starting in a world with no inequity, and a given process is the only source of inequity in subsequent years, then any inequity we observe in later years can be attributed to that process. In the preliminary simulations presented in this paper, we run each simulation for 10 years and measure the TQG at the end of each year of the simulation. The changes to the TQGs tend to level off (i.e., reach equilibrium) by 10 years, so we use the resulting TQG after 10 years in each simulation as an estimate of the amount of the overall TQG that can be explained by a given process.

6. Results

6.1. *Magnitude of Processes in Contributing to TQGs*

We report the results from the simulations described in Section 5.2 in Table 1. The simulations reported in Panel A use the proportions calculated from the North Carolina data (i.e., the left graphs in Figures 1–8), while the simulations reported in Panel B use the proportions calculated from the Washington data (i.e., the right graphs in Figures 1–8). The rows report the actual total TQG (i.e., the average difference in the proportion of low-quality teachers between disadvantaged schools and advantaged schools across the years of data used in the simulation), the simulated TQG after 10 years of the simulation (i.e., the sum of the four simulated TQGs corresponding with each process), and the share of this simulated TQG due to the different teacher labor market processes. The columns define the measure of teacher quality (experience, licensure test scores, value added without class controls, and value added with class controls), as well as the measure of school disadvantage (EDS or URM) that we use in each simulation.

The first broad takeaway is that our simulations provide reasonable though not exact approximations of the true TQGs between advantaged and disadvantaged schools. In future work, we plan to incorporate additional parameters in our simulations (e.g., the increasing number of teachers with licensure test scores from year to year or the movement of teachers between quartiles of value added from year to year) that may explain the discrepancies between our simulated TQGs and the true TQGs we observe in the data. For now, we simply note that our simulations do not appear to consistently over- or underestimate the true TQGs, which suggests that they are not systematically biased in one direction or another.

Figure 9 provides a sense of the evolution of TQGs that are calculated to be the result of the different processes in each state (Panel A for North Carolina and Panel B for Washington).

Specifically, this figure shows what the simulation suggests about the evolution in the TQG as measured by the proportion of novice teachers in high-EDS and low-EDS schools in each state. The red shaded area in Figure 9 tracks how the TQG changes in each year in the simulations in which teacher attrition is the only source of inequity. After 10 years of the simulation, this TQG is 0.026 in North Carolina (i.e., the proportion of novice teachers in disadvantaged schools is 2.6 percentage points higher than in advantaged schools) and 0.018 in Washington. These are the estimates that are reported as the “Attrition Share” in column 1 of Table 1. The remaining shaded areas of Figure 9—which we have stacked on top of each other but actually come from different simulations—represent the evolution of the TQG in the simulations in which within-district mobility (green), cross-district mobility (orange), and teacher hiring (blue) are the only sources of inequity. The share of the overall TQG explained by each process after 10 years of the simulation is then reported in column 1 of Table 1.

Now that we have walked through one set of simulations in detail, we step back to note some general patterns in Table 1. First, it is worth noting that the patterns reported are largely unaffected by the measure of student disadvantage we consider; in other words, a similar story comes through regardless of how we define school disadvantage. Second, nearly every single estimate in Table 1 is non-negative (and all but one are positive), meaning that for nearly every combination of state, measure of teacher quality, and measure of student disadvantage, our simulations show that each of the four processes contribute to larger TQGs (even if only modestly in some cases). It is not surprising, therefore, that we previously found significant TQGs across both states, all measures of teacher quality, and all measures of school disadvantage (Goldhaber et al., 2018).

Third, the definition of teacher quality does appear to matter substantially in terms of the processes that contribute most to TQGs. When we consider TQGs by teacher experience (columns 1 and 2), we see that teacher attrition, within-district mobility, and teacher hiring all contribute substantially to TQGs. In percentage terms, teacher attrition appears to explain about a third of the simulated TQG in each state, within-district mobility explains another third of the simulated TQG in North Carolina and about 10–20% of the simulated TQG in Washington, while teacher hiring explains about 20% of the simulated TQG in North Carolina and about a third of the simulated TQG in Washington. However, when we consider TQGs by teacher licensure test scores (columns 3 and 4) and value added (columns 5–8), teacher hiring explains the largest percentage of the simulated TQG in every simulation; in fact, teacher hiring explains about two thirds of the simulated TQG according to teacher licensure test scores in Washington and about three quarters of the simulated TQG according to teacher value added.

The findings on teacher hiring, particularly in Washington, suggest that there may be systematic sorting in the teacher hiring process that explains these gaps. With teacher licensure tests, for example, the teacher education programs in the state that tend to have graduates with lower licensure test scores also tend to be in the more disadvantaged parts of the state, so given the locality of the teacher labor market (e.g., Boyd, Lankford, Loeb, & Wyckoff, 2005a, 2005b; Krieg, Theobald, & Goldhaber, 2016; Reininger, 2012), these programs likely disproportionately send their graduates to disadvantaged schools. And the fact that teacher attrition explains very little of the simulated TQGs in terms of value added in either state reflects the literature (discussed in Section 2) demonstrating that differential attrition by teacher value added (i.e., ineffective teachers being particularly likely to leave disadvantaged schools) actually helps blunt the impact of the overall higher rates of attrition from disadvantaged schools on TQGs. Finally,

the importance of teacher hiring in TQGs according to value added may reflect a “dance of the lemons” in which disadvantaged schools consistently need to hire ineffective teachers from other schools when they have openings.

Finally, we do see some important differences in the simulation results between the two states. Specifically, teacher attrition and (especially) within-district mobility generally explain a larger portion of the observed TQGs in North Carolina than in Washington, while teacher hiring explains a larger portion of the TQGs in Washington than in North Carolina. The within-district mobility differences reflect differences between the two states, documented in Goldhaber et al. (2018), in terms of the distribution of TQGs across and within school districts; the majority of TQGs in North Carolina can be explained by student and teacher sorting *within* districts, while the majority of TQGs in Washington can be explained by student teacher sorting *across* districts. The differences between the states in terms of the importance of within-district mobility are not surprising given that the average district in North Carolina is about four times larger than the average district in Washington, but further work is required to understand why patterns in teacher hiring appear to explain so much of the observed inequity in Washington.

6.2. Persistence of Processes Contributing to TQGs

The simulations discussed in Section 6.1 pool all proportions across the years of available data in each state. However, it is potentially important to understand the extent to which these proportions vary across these years of data. Specifically, we hypothesize that processes for which inequities are more persistent across the years of data may be more difficult to change. In other words, a process that is consistently inequitable from year to year may be the product of entrenched policies or structures that are difficult to change, while a process that is much more equitable in some years than others may be more malleable to policy interventions.

To explore the persistence in the inequity in each of the processes considered in this paper, we first calculate the differences in attrition rates, within-district transfer rates, cross-district transfer rates, and hiring rates of ineffective teachers between advantaged and disadvantaged schools for each year of data in each state. Then across the years of data in each state, we calculate the coefficient of variation of these differences for each process (i.e., the standard deviation of the differences divided by the mean of the differences). This statistic provides preliminary estimates of the persistence of the inequities in each process.

Our overall conclusion from this exercise is that the inequities in these processes are similarly persistent, with a few important exceptions. In North Carolina, inequities in attrition and hiring by licensure test scores are considerably more persistent (i.e., have lower coefficients of variation across years) than inequities in the other processes, while in Washington, only the inequities in hiring by licensure test scores are notably more persistent; in fact, while the coefficient of variation for hiring by licensure test scores is 0.14 in Washington, all the other coefficients of variation are between 0.6 and 0.8. This provides an interesting addendum to the simulation results in Section 6.1; namely, while inequities in teacher hiring are clearly quite important in terms of their contribution to TQGs by licensure test scores, these inequities may be harder to address because they have been extremely persistent over time.¹⁹ This is also consistent with the hypothesis, discussed in Section 6.1, that inequities in teacher hiring could be the result of structural factors, like the sorting of teachers from teacher education programs to nearby districts, that may be difficult to change.

¹⁹ One could also argue, though, that this is an area of potential improvement given that schools do not typically have access to teacher licensure test scores (i.e., providing access to these scores could allow schools to address this inequity in a way they could not before).

7. Conclusions

As discussed in the introduction, the simulations described in this paper make a number of simplifying assumptions and therefore do not represent a true dynamic model of teacher attrition, mobility, and hiring (i.e., we may not want to use the stochastic model in this paper to *project* how TQGs might change in response to a policy intervention).²⁰ However, we still believe this exercise is useful because it is well suited for explaining why the TQGs we observe in public schools exist in the first place (as evidenced by the fact that they provide reasonable approximations of the observed TQGs in the years of data we consider). Specifically, the results discussed in Section 6 provide some of the first empirical evidence about the extent to which teacher attrition, mobility, and hiring contribute to TQGs between advantaged and disadvantaged schools.

We also believe that these results can inform some preliminary recommendations for policymakers seeking to close TQGs, though we caution that these simulations provide an answer to a very narrow question: about how much of the overall TQG can we expect to be removed if we made one of these processes completely equitable between advantaged and disadvantaged schools? This is *not* the only relevant question for policy purposes; issues like the cost-efficiency of various interventions should surely play a large role in decisions about closing TQGs. Moreover, it may be even more efficient to close TQGs by making a process inequitable (in favor of disadvantaged schools) to offset inequities in another process that may be harder to address. That said, we believe that it may be more politically feasible to design policies that seek to create equality in these processes (e.g., to keep the same percentage of high-quality teachers in

²⁰ We do plan to develop a more complicated dynamic model that will allow us to make these types of projections.

advantaged and disadvantaged schools) and thus that the answers to the admittedly narrow question our simulations address are still relevant to policymakers.

Our most striking finding is the outsized role that teacher hiring plays in contributing to TQGs as measured by licensure test scores and value added, particularly given that there is little prior empirical evidence about differences in the proportion of teachers with low licensure test scores or low prior estimates of value added who are hired into advantaged and disadvantaged schools. This finding might suggest that policymakers should consider recruitment or hiring policies to attract higher-quality teachers to disadvantaged schools and close these TQGs. However, this policy recommendation is somewhat tempered by our finding that inequities in hiring by teacher licensure test scores have been extremely persistent across the years of data in both states, which suggests that these inequities may be difficult to address. The bottom line, though, is that this result certainly suggests that we need to understand much more about the teacher hiring process and *why* teacher hiring contributes so much to TQGs between advantaged and disadvantaged schools.

We also find that teacher attrition and mobility contribute substantially to TQGs as measured by teacher experience, which supports the arguments for interventions that keep high-quality teachers in disadvantaged schools (or just reduce overall levels of teacher attrition from disadvantaged schools) that are often the focus of policy discussions (e.g., Carver-Thomas & Darling-Hammond, 2017). An oft-cited example of such an intervention, in one of the focal states considered in this study, is a bonus policy implemented in North Carolina in the early 2000s in which certified math, science, and special education teachers working in disadvantaged schools (either with high poverty rates or low test scores) received an annual bonus of \$1,800. Clotfelter et al. (2008) found strong evidence that the policy reduced the attrition of teachers

from disadvantaged schools and, importantly, found that experienced teachers demonstrated the greatest response to the policy. Our simulations suggest that such an intervention may have a substantial impact on TQGs in terms of teacher experience in the state given the importance of teacher attrition to these TQGs. In future work, we plan to incorporate point estimates from studies like Clotfelter et al. (2008) into the simulation framework introduced in this paper to estimate the potential impact of similar policies on overall TQGs.

References

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25, 95–135.
- Barnes, G., Crowe, E., & Schaefer, B. (2007). The Cost of Teacher Turnover in Five School Districts: A Pilot Study. *National Commission on Teaching and America's Future*.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Who leaves? Teacher attrition and student achievement* (NBER Working Paper No. 14022). Cambridge, MA: National Bureau of Economic Research.
- Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short term careers of high-achieving teachers in schools with low performing students. *American economic review*, 95(3), 166-171
- Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005b). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, 24, 113–132.
- Carver-Thomas, D., & Darling-Hammond, L. (2017). *Teacher turnover: Why it matters and what we can do about it*. Palo Alto, CA: Learning Policy Institute.
- Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, 92, 1352–1370.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, 24, 377–392.
- Clotfelter, C., Ladd, H. F., Vigdor, J., & Wheeler, J. (2007). High poverty schools and the distribution of teachers and principals. *North Carolina Law Review*, 85, 1345–1379.
- Cook, J. (2011). *Within-school teacher quality: A dynamic model of teacher mobility*. Chapel Hill, NC: University of North Carolina at Chapel Hill.
- Feng, L., & Sass, T. R. (2017). Teacher quality and teacher mobility. *Education Finance and Policy*, 12, 396–418.
- Goldhaber, D., Gratz, T., & Theobald, R. (2017). What's in a teacher test? Assessing the relationship between teacher licensure test scores and student secondary STEM achievement and course taking. *Economics of Education Review*, 61, 112–129.
- Goldhaber, D., Gross, B., & Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management*, 30, 57–87.

- Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher, 44*, 293–307.
- Goldhaber, D., Quince, V., & Theobald, R. (2016). *Reconciling different estimates of teacher quality gaps based on value-added*. Washington, DC: National Center for the Analysis of Longitudinal Data in Education Research.
- Goldhaber, D., Quince, V., & Theobald, R. (2018). Has it always been this way? Tracing the evolution of teacher quality gaps in U.S. public schools. *American Educational Research Journal, 55*, 171–201. doi 10.3102/0002831217733445
- Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record, 113*, 2552–2585.
- Gritz, M., & Theobald, N. (1996). The effects of school district spending priorities on length of stay in teaching. *Journal of Human Resources, 31*, 477–512.
- Hanushek, E. A., & Rivkin, S. G. (2010). *Constrained job matching: Does teacher job search harm disadvantaged urban schools?* (No. w15816). National Bureau of Economic Research.
- Hanushek, E., Kain, J., & Rivken, S. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*, 326–354.
- Imazeki, J. (2005). Teacher salaries and teacher attrition. *Economics of Education Review, 24*, 431–449.
- Ingersoll, R. M., & May, H. (2012). The magnitude, destinations, and determinants of mathematics and science teacher turnover. *Educational Evaluation and Policy Analysis, 34*, 435–464.
- Isenberg, E., Max, J., Gleason, P., Johnson, M., Deutsch, J., & Hansen, M. (2016). Do low-income students have equal access to effective teachers? Evidence from 26 districts (NCEE 2017-4007). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- 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*(4), 85–108.

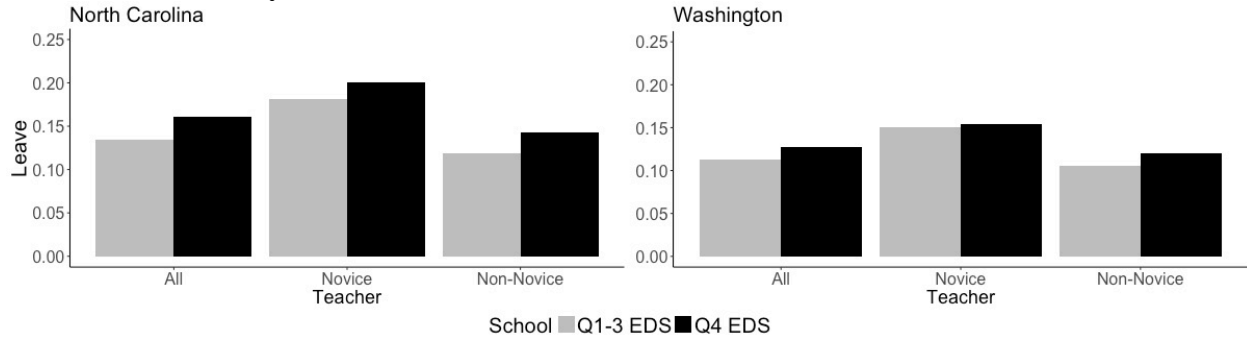
- Kaiser, A. (2011). Beginning teacher attrition and mobility: Results from the first through third waves of the 2007–08 Beginning Teacher Longitudinal Study. First look (NCES 2011-318). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude and effects of student sorting within schools. *Educational Researcher*, 42, 304–316.
- Keigher, A. (2010). Teacher attrition and mobility: Results from the 2008–09 Teacher Follow-up Survey. First look (NCES 2010-353). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- Krieg, J. M. (2006). Teacher quality and attrition. *Economics of Education Review*, 25, 13–27.
- Krieg, J. M., Theobald, R., & Goldhaber, D. (2016). A foot in the door: Exploring the role of student teaching assignments in teachers' initial job placements. *Educational Evaluation and Policy Analysis*, 38, 364–388.
- Kyle Ingle, W. (2009). Teacher quality and attrition in a US school district. *Journal of Educational Administration*, 47(5), 557-585.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24, 37–62.
- Loeb, S., Kalogrides, D., & Bételle, T. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy*, 7, 269–304.
- Marvel, J., Lyter, D. M., Peltola, P., Strizek, G. A., Morton, B. A., & Rowland, R. (2007). Teacher attrition and mobility: Results from the 2004–05 Teacher Follow-up Survey (NCES 2007-307). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education*, 4, 572–606.
- Reform Support Network. (2015, March). *Promoting more equitable access to effective teachers: Problem and root causes*. Washington, DC: U.S. Department of Education.
- 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, 127–145.
- Rice, J. K. (2013). Learning from experience? Evidence on the impact and distribution of teacher experience and the implications for teacher policy. *Education*, 8, 332–348.

- Rich, M. (2014, November 10). U.S. to focus on equity in assigning of teachers. *New York Times*, A18.
- Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94, 247–252.
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4-36.
- Sass, T., Hannaway, J., Xu, Z., Figlio, D., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower-poverty schools. *Journal of Urban Economics*, 72, 104–122.
- Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review*, 26, 145–159.
- Shen, J. (1997). Teacher retention and attrition in public schools: Evidence from SASS91. *Journal of Educational Research*, 91, 81–88.
- Xu, Z., Ozek, U., & Hansen, M. (2015). Teacher performance trajectories in high- and lower-poverty schools. *Educational Evaluation and Policy Analysis*, 37, 459–477.

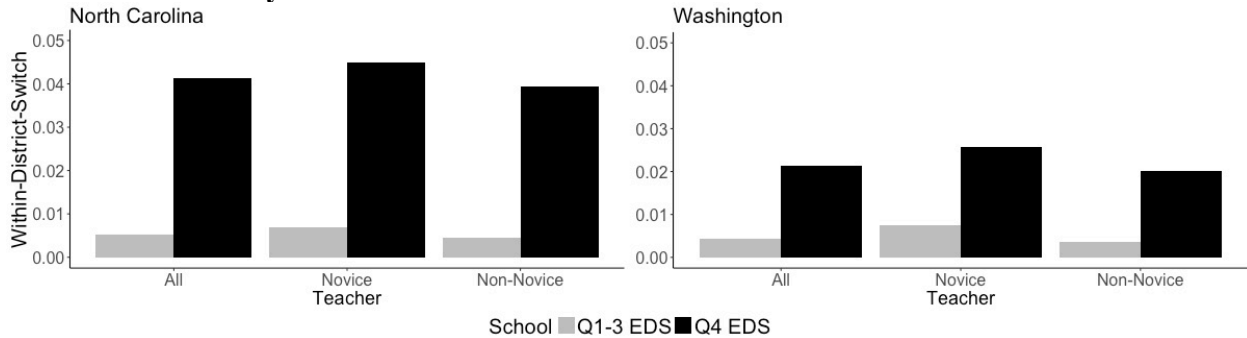
Figures and Tables

Figure 1. Teacher mobility and hiring by school EDS and teacher experience

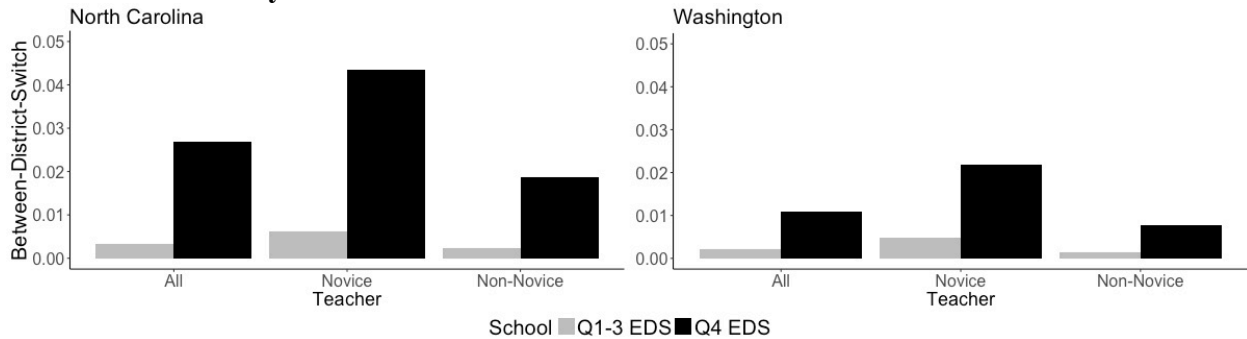
Panel A. Probability of teacher attrition



Panel B. Probability of teacher within-district move to different level of school EDS



Panel C. Probability of teacher between-district move to different level of school EDS



Panel D. Probability of teacher hiring into open positions of school EDS

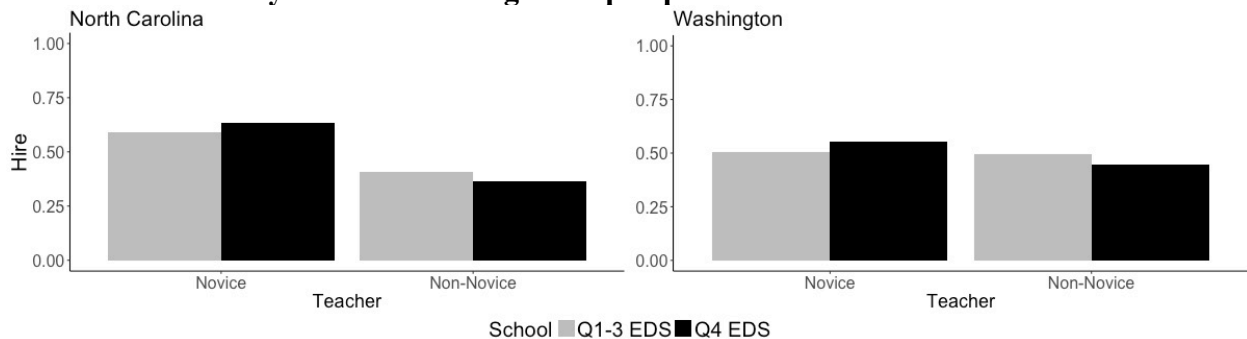
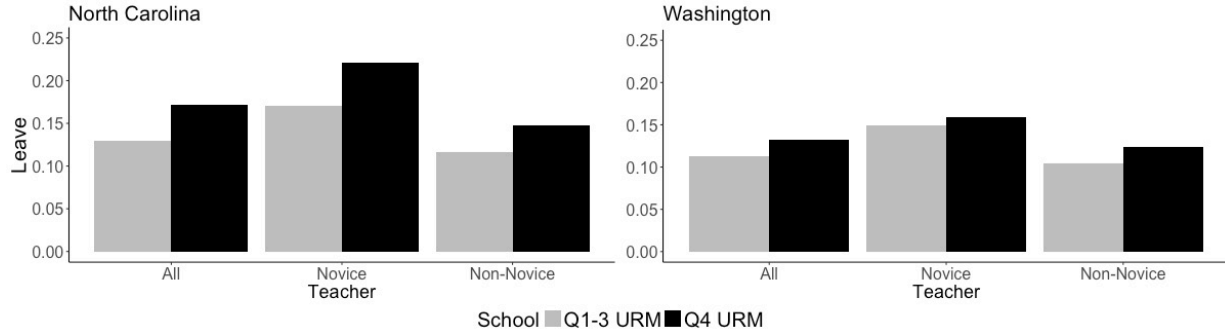
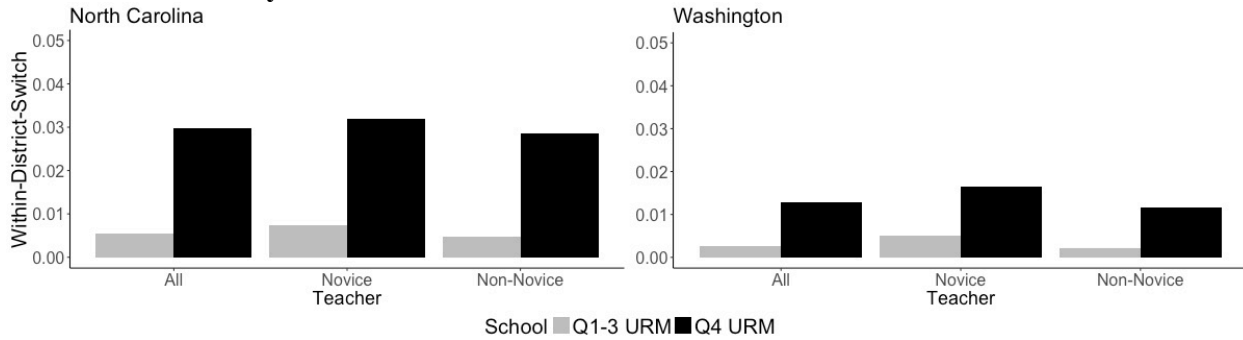


Figure 2. Teacher mobility and hiring by school URM and teacher experience

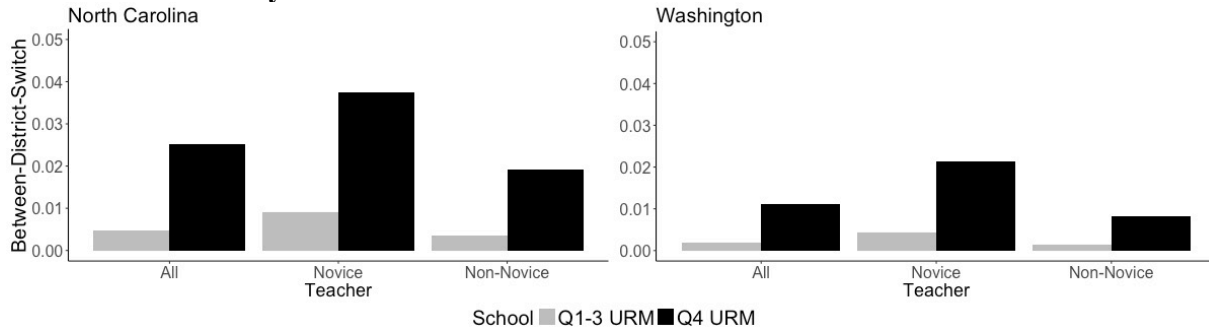
Panel A. Probability of teacher attrition by school URM



Panel B. Probability of teacher within-district move to different level of school URM



Panel C. Probability of teacher between-district move to different level of school URM



Panel D. Probability of teacher hiring into open positions by school URM

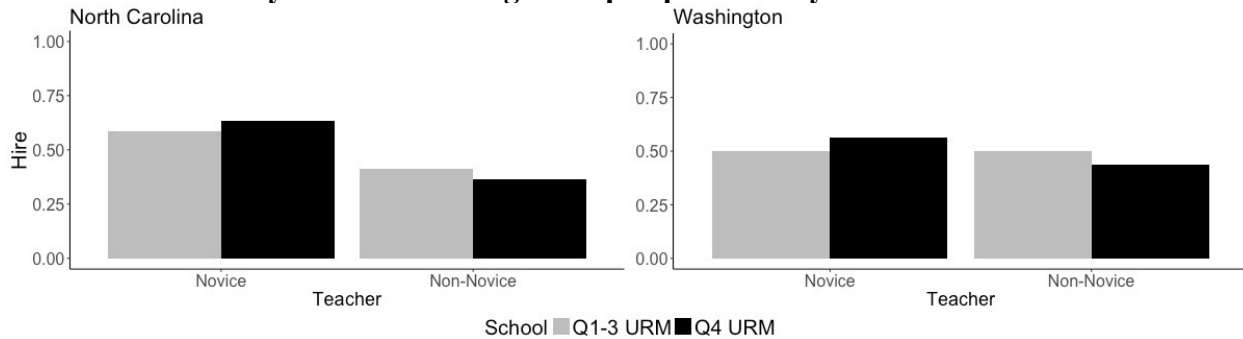
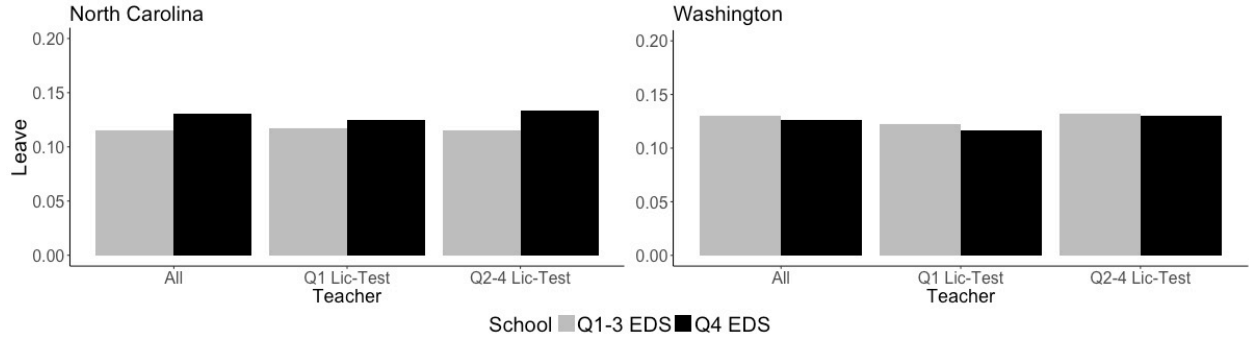
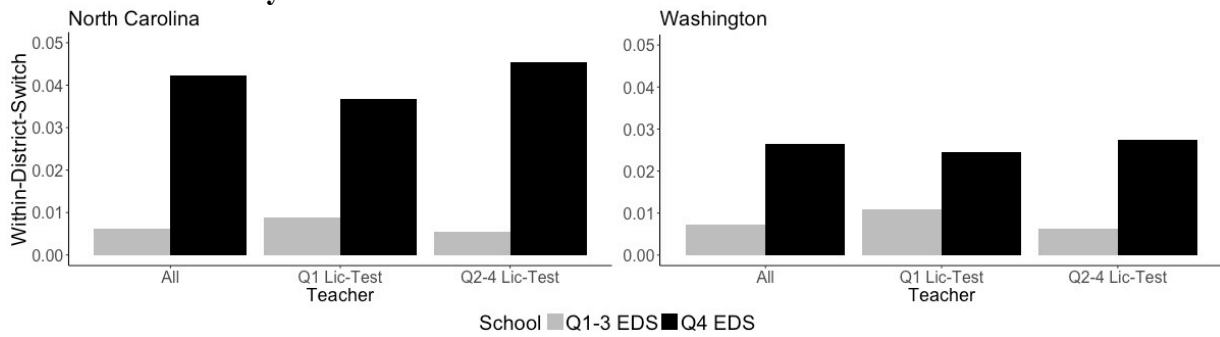


Figure 3. Teacher mobility and hiring by school EDS and teacher licensure test scores

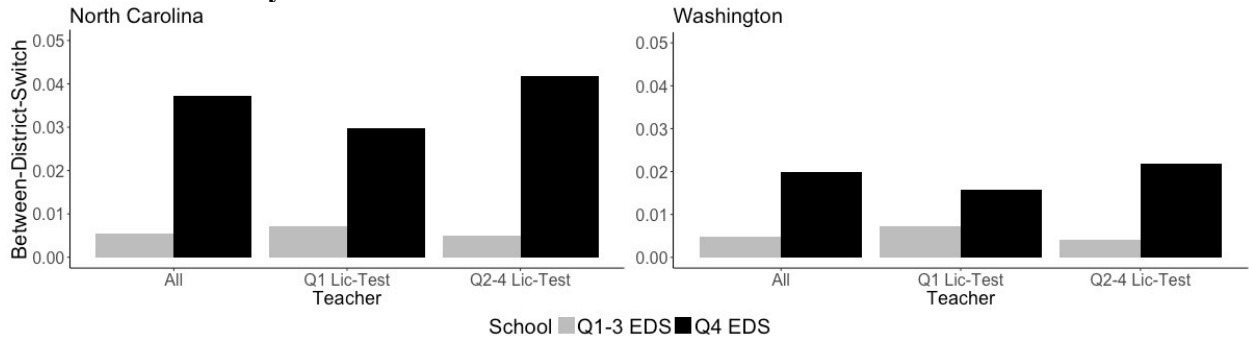
Panel A. Probability of teacher attrition



Panel B. Probability of teacher within-district move to different level of school EDS



Panel C. Probability of teacher between-district move to different level of school EDS



Panel D. Probability of teacher hiring into open positions

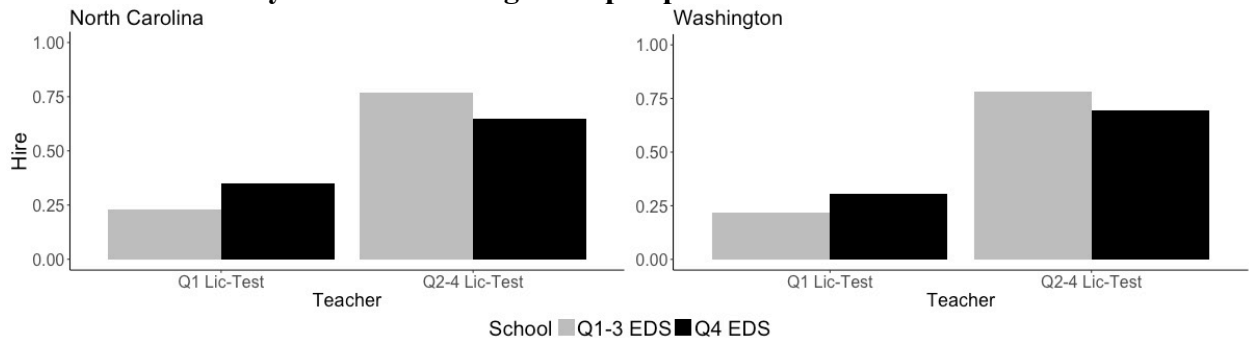
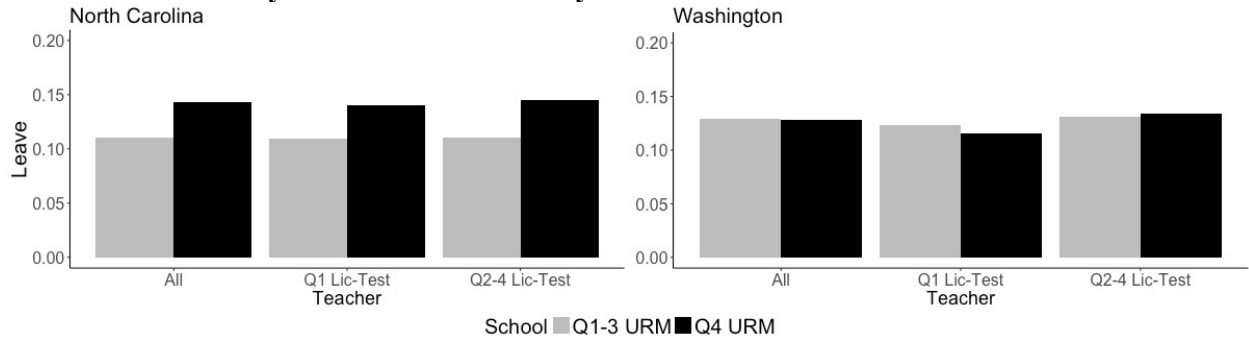
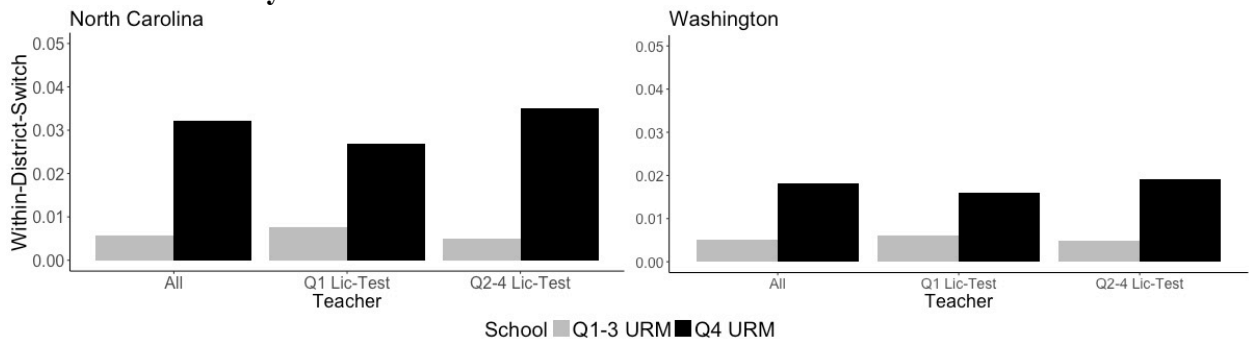


Figure 4. Teacher mobility and hiring by school URM and teacher licensure test scores

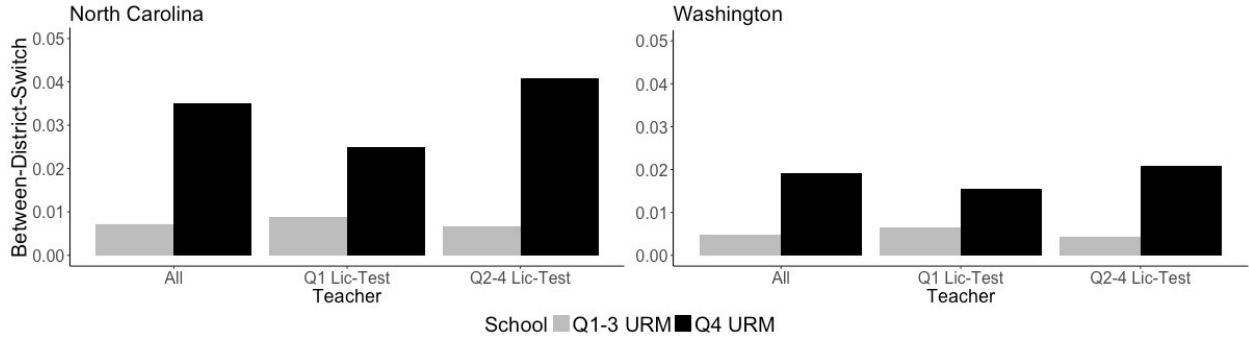
Panel A. Probability of teacher attrition by school URM



Panel B. Probability of teacher within-district move to different level of school URM



Panel C. Probability of teacher between-district move to different level of school URM



Panel D. Probability of teacher hiring into open positions by school URM

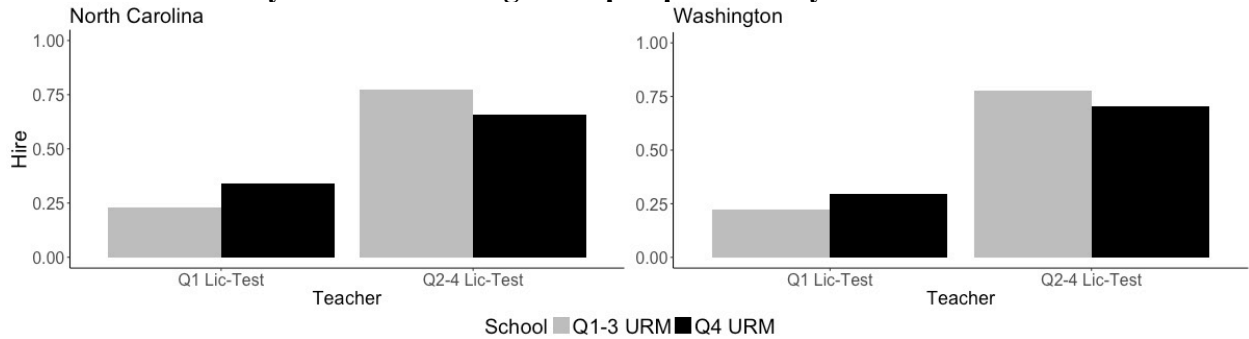
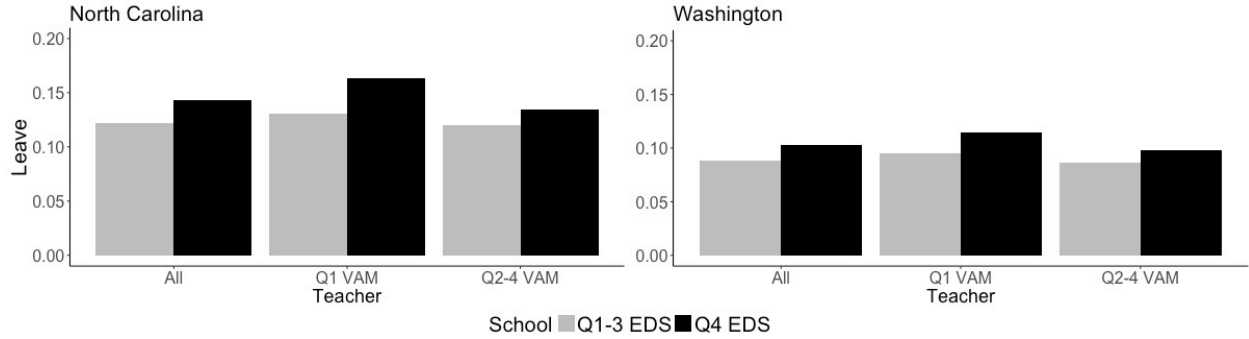
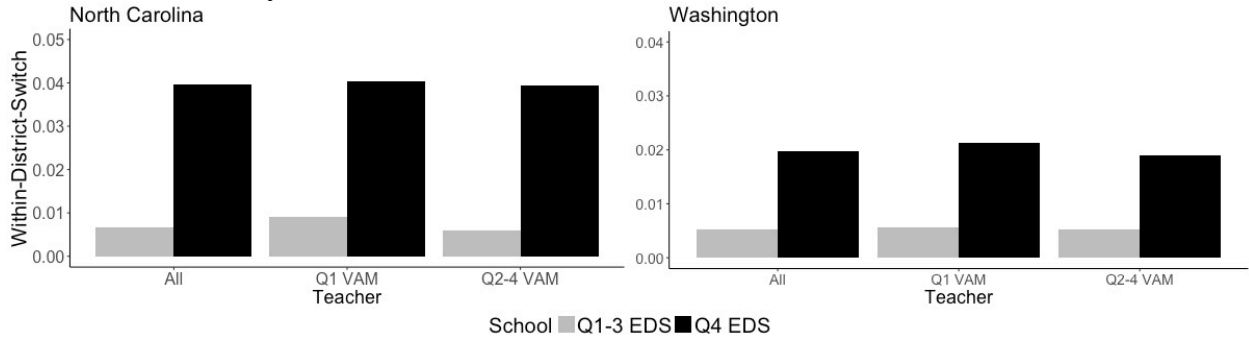


Figure 5. Teacher mobility and hiring by school EDS and teacher value added

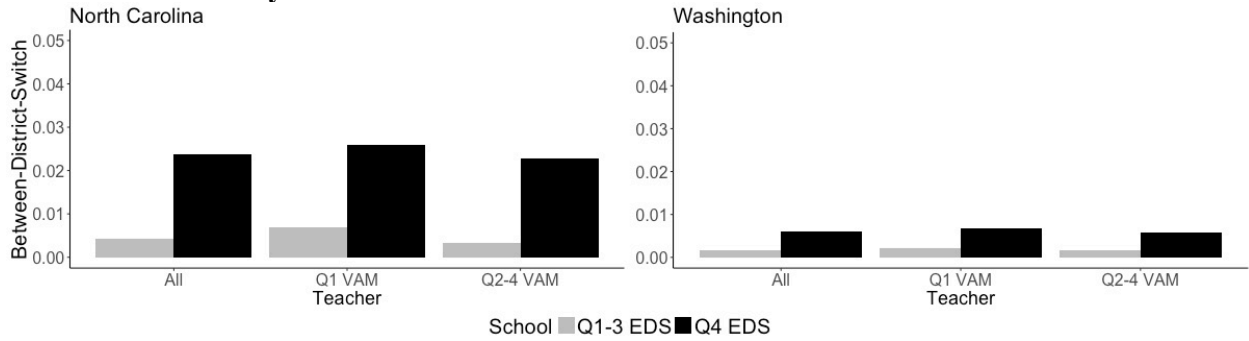
Panel A. Probability of teacher attrition



Panel B. Probability of teacher within-district move to different level of school EDS



Panel C. Probability of teacher between-district move to different level of school EDS



Panel D. Probability of teacher hiring into open positions

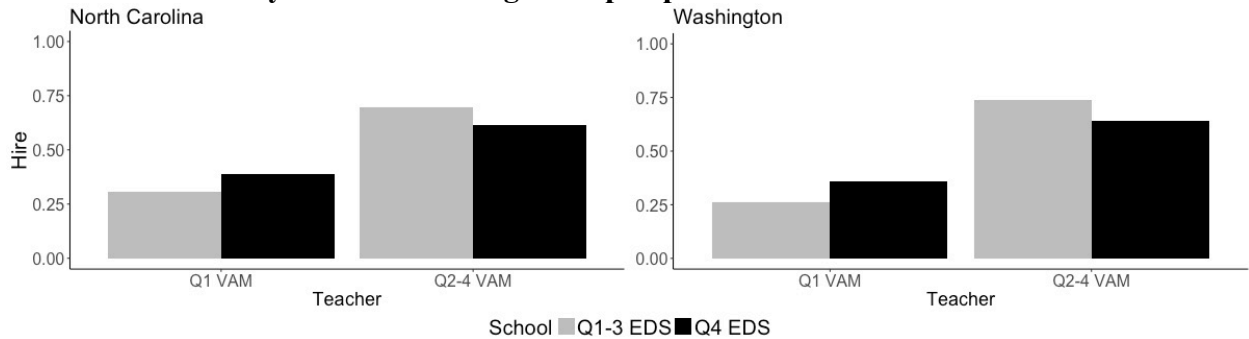
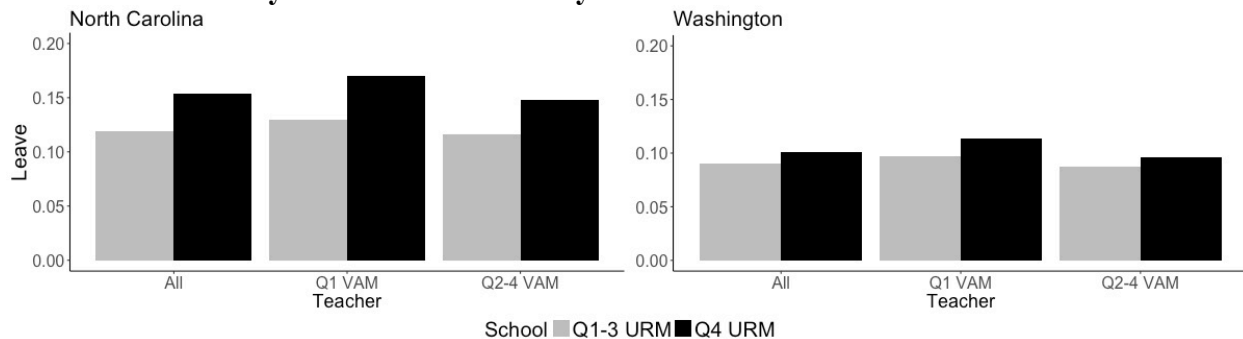
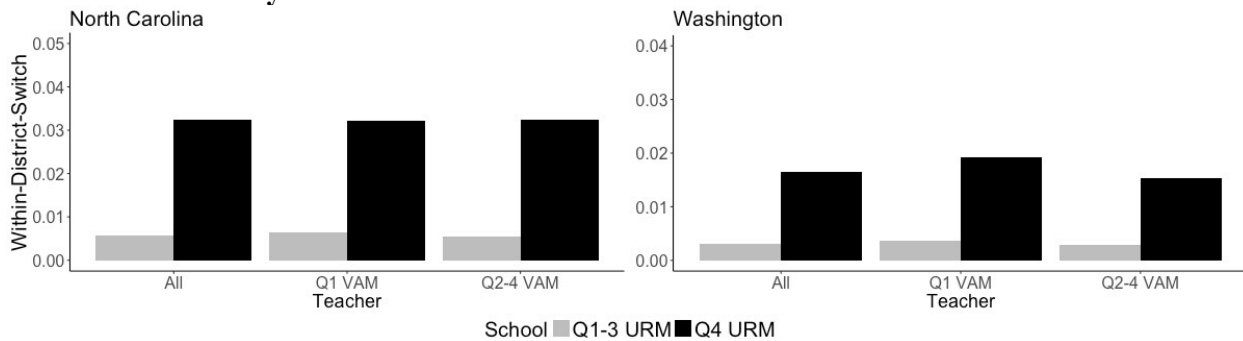


Figure 6. Teacher mobility and hiring by school URM and teacher value added

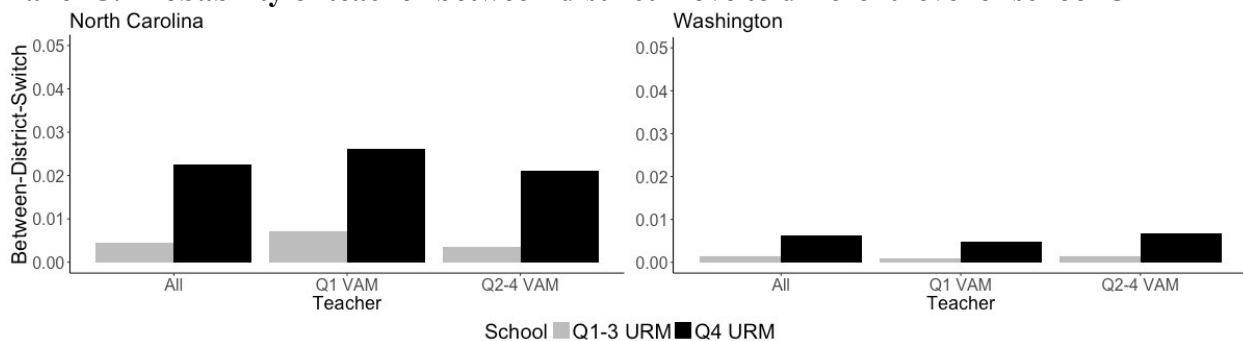
Panel A. Probability of teacher attrition by school URM



Panel B. Probability of teacher within-district move to different level of school URM



Panel C. Probability of teacher between-district move to different level of school URM



Panel D. Probability of teacher hiring into open positions

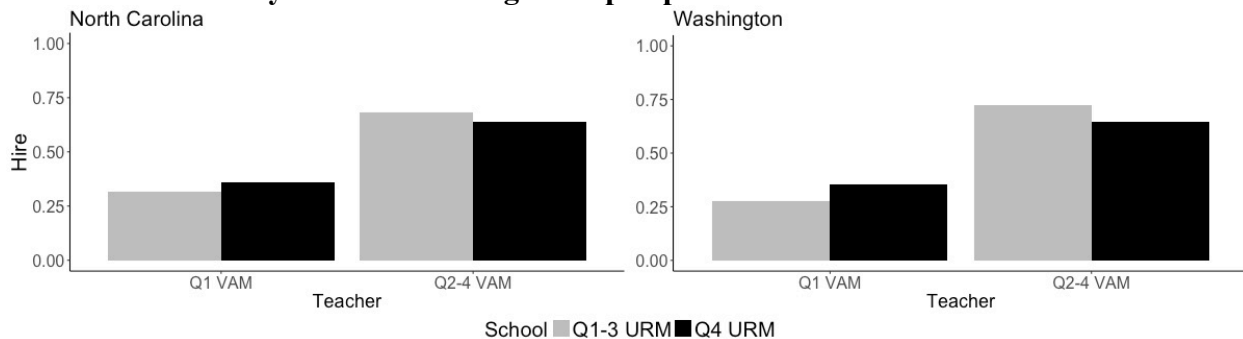
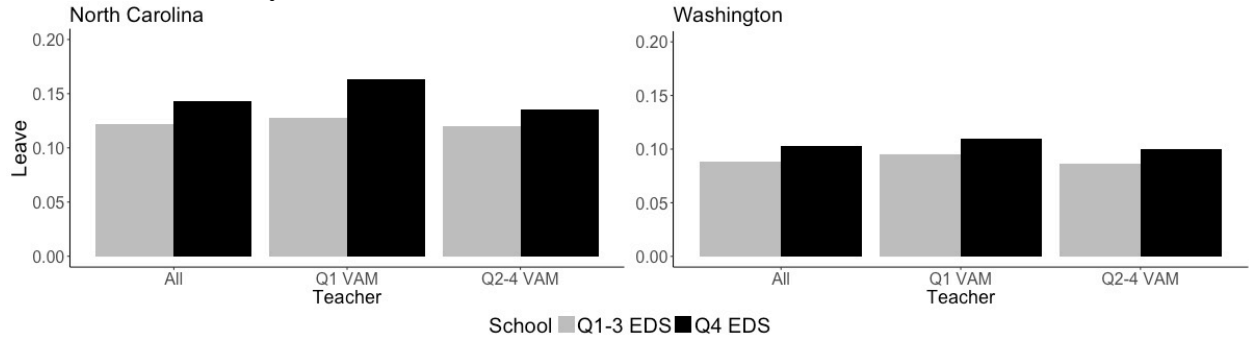
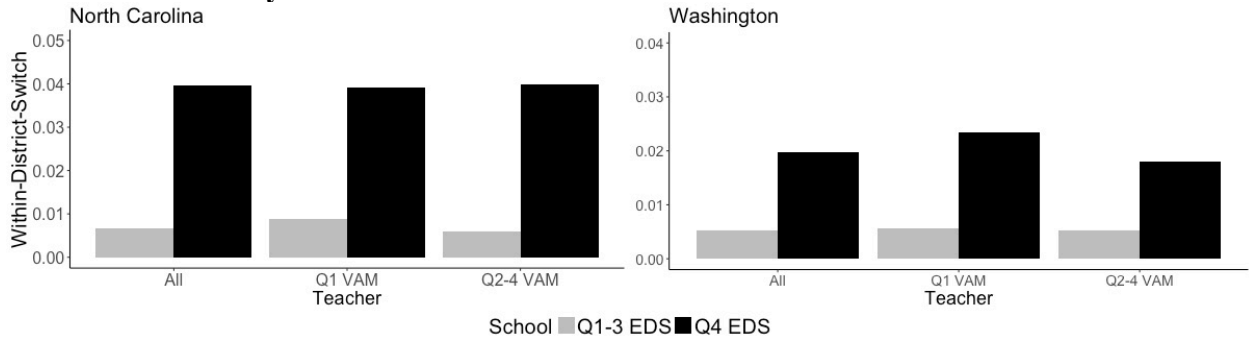


Figure 7. Teacher mobility and hiring by school EDS and teacher value added (classroom controls)

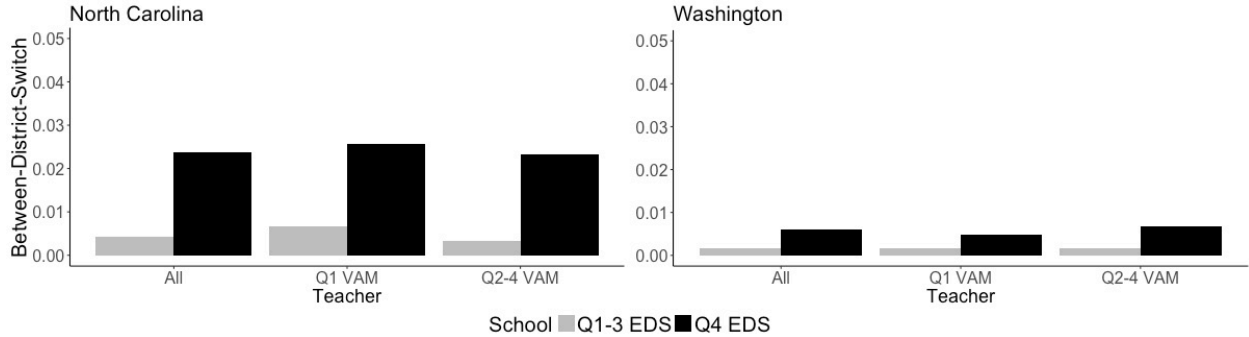
Panel A. Probability of teacher attrition



Panel B. Probability of teacher within-district move to different level of school EDS



Panel C. Probability of teacher between-district move to different level of school EDS



Panel D. Probability of teacher hiring into open positions

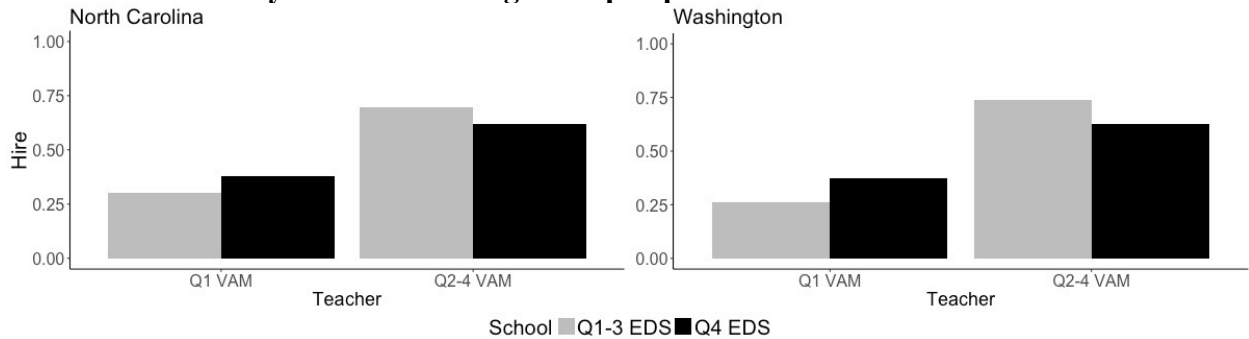
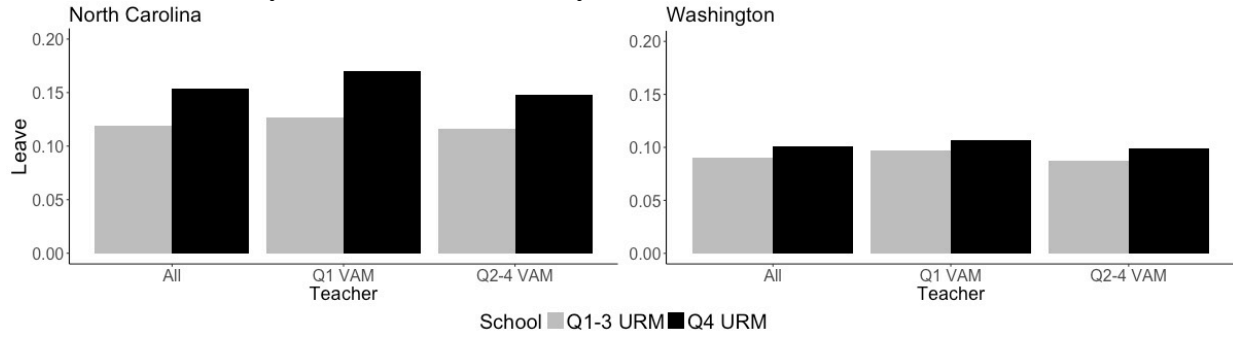
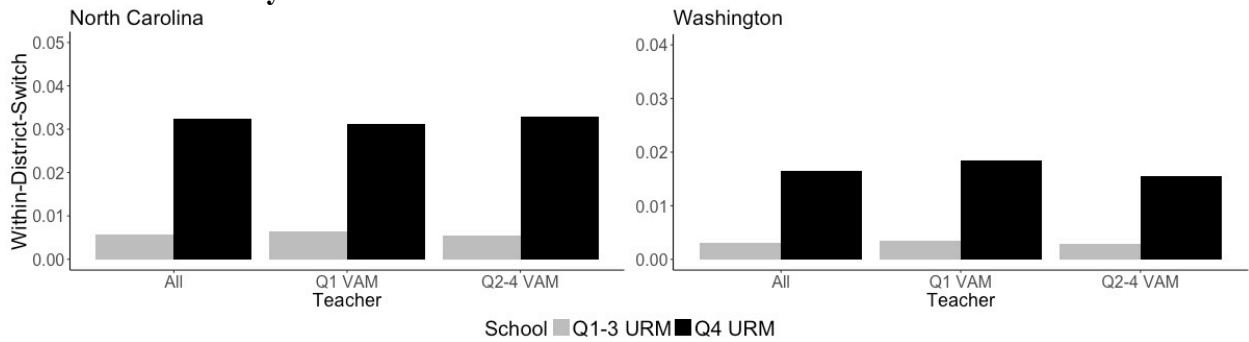


Figure 8. Teacher mobility and hiring by school URM and teacher value added (classroom controls)

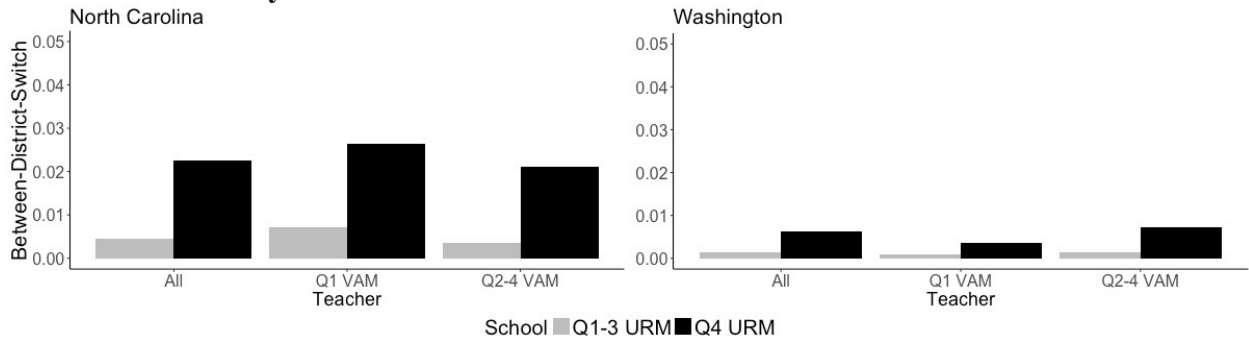
Panel A. Probability of teacher attrition by school URM



Panel B. Probability of teacher within-district move to different level of school URM



Panel C. Probability of teacher between-district move to different level of school URM



Panel D. Probability of teacher hiring into open positions

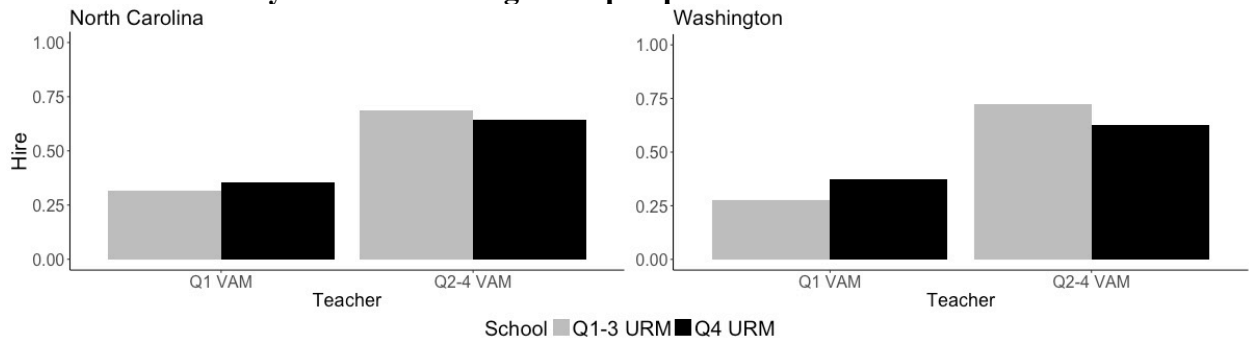
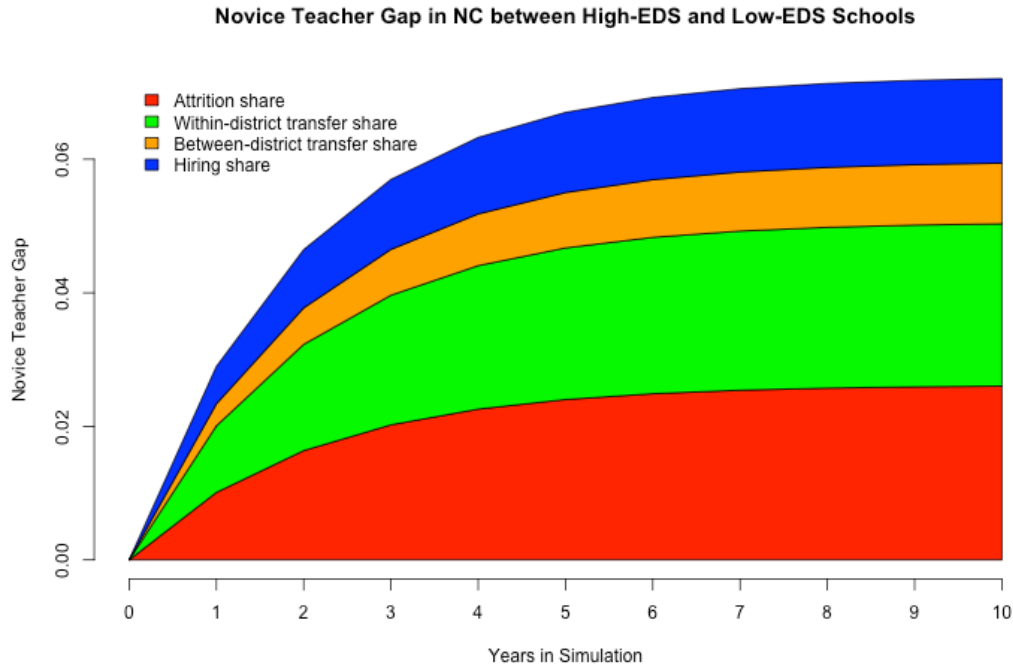


Figure 9. Simulation results by school EDS and teacher experience
Panel A. Simulation results in North Carolina



Panel B. Simulation results in Washington

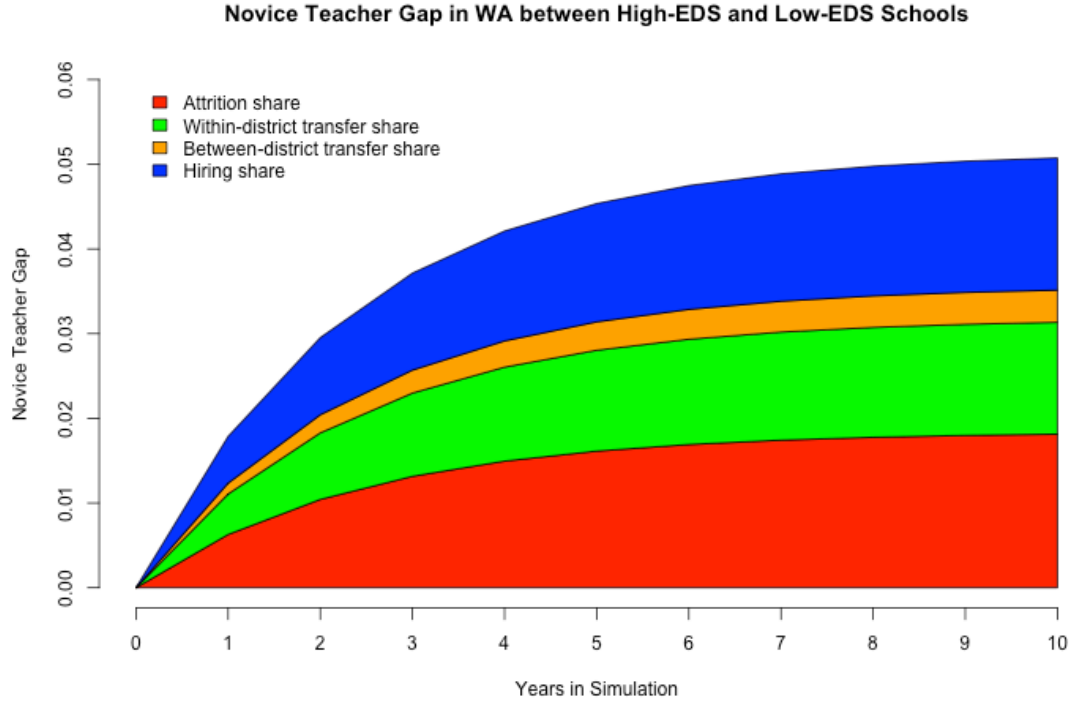


Table 1. Simulation results after 10 years

Measure of Teacher Quality:	Experience		Licensure Tests		Value Added (No Class Controls)		Value Added (Class Controls)	
Measure of Student Disadvantage:	EDS	URM	EDS	URM	EDS	URM	EDS	URM
Column:	1	2	3	4	5	6	7	8
Panel A: North Carolina								
Actual Gap	0.072	0.083	0.145	0.132	0.061	0.041	0.040	0.023
Simulated Gap	0.071	0.083	0.146	0.132	0.060	0.052	0.058	0.046
Attrition Share	0.026	0.033	0.014	0.011	0.007	0.000	0.007	-0.004
Within-District Mobility Share	0.024	0.020	0.027	0.022	0.018	0.009	0.018	0.010
Cross-District Mobility Share	0.009	0.011	0.026	0.028	0.011	0.012	0.011	0.011
Hiring Into Open Positions Share	0.012	0.019	0.079	0.071	0.024	0.031	0.022	0.029
Panel B: Washington								
Actual Gap	0.048	0.064	0.080	0.068	0.063	0.068	0.059	0.070
Simulated Gap	0.051	0.067	0.069	0.070	0.051	0.039	0.064	0.058
Attrition Share	0.018	0.026	0.000	0.007	0.003	0.000	0.009	0.008
Within-District Mobility Share	0.013	0.008	0.013	0.007	0.005	0.004	0.002	0.004
Cross-District Mobility Share	0.004	0.004	0.014	0.013	0.002	0.002	0.004	0.004
Hiring Into Open Positions Share	0.016	0.029	0.042	0.043	0.041	0.033	0.049	0.042

Appendix: Simulation Details

For each combination of state, teacher quality measure, and measure of school disadvantage, we perform four simulations corresponding with the four conditions described in the main text:

1. Teacher attrition is the only source of inequity
2. Within-district teacher mobility is the only source of inequity
3. Cross-district teacher mobility is the only source of inequity
4. Teacher hiring is the only source of inequity

As described in the main text, each simulation begins with starting values such that $p_{LQ,t}^{DIS} = p_{LQ,t}^{ADV}$ (i.e., there is no teacher quality gap in year t). To describe the remainder of the simulation under each condition, we define *non-school-type-specific* definitions of the proportions described in the previous section:

- $p_{Q,t}^{WD}$ = the proportion of teachers of quality Q that move within a district after year t
- $p_{Q,t}^{BD}$ = the proportion of teachers of quality Q that move between districts after year t
- $p_{Q,t}^{ATT}$ = the proportion of teachers of quality Q that leave the workforce after year t
- $p_{Q,(t+1)}^{HIRE}$ = the proportion of teachers of quality Q hired into open positions in year $t + 1$

We describe below how we use these proportions in the simulation under each condition.

1 Teacher attrition is the only source of inequity

In this simulation, we set $p_{Q,t}^{ATT,DIS}$ and $p_{Q,t}^{ATT,ADV}$ to their observed values for $Q \in \{LQ, HQ\}$, but set all other proportions in the simulation to non-school-type-specific values:

$$\begin{aligned}
 p_{Q,t}^{WD,ADV \rightarrow DIS} &= p_{Q,t}^{WD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{WD,ADV \rightarrow ADV} &= p_{Q,t}^{WD,DIS \rightarrow DIS} &= p_{Q,t}^{WD} \\
 p_{Q,t}^{BD,ADV \rightarrow DIS} &= p_{Q,t}^{BD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{BD,ADV \rightarrow ADV} &= p_{Q,t}^{BD,DIS \rightarrow DIS} &= p_{Q,t}^{BD} \\
 p_{Q,(t+1)}^{HIRE,ADV} &= p_{Q,(t+1)}^{HIRE,DIS} &= p_{Q,(t+1)}^{HIRE}
 \end{aligned}$$

2 Within-district teacher mobility is the only source of inequity

In this simulation, we set $p_{Q,t}^{WD,ADV \rightarrow DIS}$, $p_{Q,t}^{WD,DIS \rightarrow ADV}$, $p_{Q,t}^{WD,ADV \rightarrow ADV}$, and $p_{Q,t}^{WD,DIS \rightarrow DIS}$ to their observed values for $Q \in \{LQ, HQ\}$, but set all other proportions in the simulation to non-school-type-specific values:

$$\begin{aligned}
 p_{Q,t}^{ATT,DIS} &= p_{Q,t}^{ATT,ADV} &= p_{Q,t}^{ATT} \\
 p_{Q,t}^{BD,ADV \rightarrow DIS} &= p_{Q,t}^{BD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{BD,ADV \rightarrow ADV} &= p_{Q,t}^{BD,DIS \rightarrow DIS} &= p_{Q,t}^{BD} \\
 p_{Q,(t+1)}^{HIRE,ADV} &= p_{Q,(t+1)}^{HIRE,DIS} &= p_{Q,(t+1)}^{HIRE}
 \end{aligned}$$

3 Cross-district teacher mobility is the only source of inequity

In this simulation, we set $p_{Q,t}^{BD,ADV \rightarrow DIS}$, $p_{Q,t}^{BD,DIS \rightarrow ADV}$, $p_{Q,t}^{BD,ADV \rightarrow ADV}$, and $p_{Q,t}^{BD,DIS \rightarrow DIS}$ to their observed values for $Q \in \{LQ, HQ\}$, but set all other proportions in the simulation to non-school-type-specific values:

$$\begin{aligned}
 p_{Q,t}^{ATT,DIS} &= p_{Q,t}^{ATT,ADV} &= p_{Q,t}^{ATT} \\
 p_{Q,t}^{WD,ADV \rightarrow DIS} &= p_{Q,t}^{WD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{WD,ADV \rightarrow ADV} &= p_{Q,t}^{WD,DIS \rightarrow DIS} &= p_{Q,t}^{WD} \\
 p_{Q,(t+1)}^{HIRE,ADV} &= p_{Q,(t+1)}^{HIRE,DIS} &= p_{Q,(t+1)}^{HIRE}
 \end{aligned}$$

4 Teacher hiring is the only source of inequity

In this simulation, we set $p_{Q,(t+1)}^{HIRE,ADV}$ and $p_{Q,(t+1)}^{HIRE,DIS}$ to their observed values for $Q \in \{LQ, HQ\}$, but set all other proportions in the simulation to non-school-type-specific values:

$$\begin{aligned}
 p_{Q,t}^{ATT,DIS} &= p_{Q,t}^{ATT,ADV} &= p_{Q,t}^{ATT} \\
 p_{Q,t}^{WD,ADV \rightarrow DIS} &= p_{Q,t}^{WD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{WD,ADV \rightarrow ADV} &= p_{Q,t}^{WD,DIS \rightarrow DIS} &= p_{Q,t}^{WD} \\
 p_{Q,t}^{BD,ADV \rightarrow DIS} &= p_{Q,t}^{BD,DIS \rightarrow ADV} &= 0 \\
 p_{Q,t}^{BD,ADV \rightarrow ADV} &= p_{Q,t}^{BD,DIS \rightarrow DIS} &= p_{Q,t}^{BD}
 \end{aligned}$$