Leadership and Policy in Schools

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/nlps20

Staffing for School Turnaround in Rural Settings

Erica Harbatkin

LEADERSHIP AND POLICY IN SCHOOLS
https://doi.org/10.1080/15700763.2022.2058963

Staffing for School Turnaround in Rural Settings

Erica Harbatkin

Michigan State University, Education Policy Innovation Collaborative, East Lansing, Michigan, USA

ABSTRACT
Recruiting and retaining effective teachers is critical to school turnaround. However, research on how improving educator quality in low-performing schools contributes to school improvement is largely situated in urban settings. This study examines staffing practices through a descriptive analysis of the first cohort of Comprehensive Support and Improvement (CSI) schools in North Carolina. In the 94 schools slated for turnaround, ineffective teachers in rural schools were less likely to turn over than in non-rural schools. Instead of filling vacancies with more effective teachers than they lost, rural schools tended to assign their more effective teachers to tested grades and subjects.

Introduction

A large literature documents the importance of effective teachers for student learning (Aaronson et al., 2007; Chetty et al., 2014; Rivkin et al., 2005; Rockoff, 2004) Research shows that school and district leaders engage in a variety of strategies to improve teacher quality in low-performing schools in urban settings (see, e.g., Adnot et al., 2017; Dee & Wyckoff, 2015; Loeb et al., 2015). A smaller but growing literature on rural schools suggests that rural teacher labor markets are less robust and rural school leaders may therefore need to employ different strategies than their urban counterparts to build and sustain a stable and high quality teaching staff (see, e.g., Nguyen, 2020). Even less is known about staffing low-performing rural schools – though recruiting, developing, and retaining a highly effective educator workforce is critical to successful school turnaround and has become even more salient as schools face staffing shortages in the wake of the COVID-19 pandemic. Teacher burnout was extraordinarily high during the pandemic, and as the 2021–22 school year began, early data showed staffing shortages were pronounced throughout the country (Barnum, 2021; Kaufman & Diliberti, 2021). These shortages were especially evident in high poverty schools and districts – the same schools that tended to be designated as low performing even before the pandemic (Goldhaber & Gratz, 2021).

Prior to the pandemic, one way low-performing schools aimed to raise teacher effectiveness was through strategic staffing practices in which they aimed to hire teachers who were more effective than the teachers who left (Adnot et al., 2017; Cullen et al., 2019; Dee & Wyckoff, 2015; James & Wyckoff, 2020; Weisberg et al., 2009). In fact, these strategies were codified in school turnaround policy under Race to the Top (RTTT) and School Improvement Grants (SIG), which in some cases called for widespread staff replacement under the assumption that the compositional effect of this staff replacement would be positive; in other words, the replacement teachers would be more effective, on average, than the teachers they replaced. But unlike schools in urban and many suburban settings, rural schools often do not have a robust pool of highly effective teachers from which to recruit replacements. Recruitment and retention efforts in low-performing rural schools are further hindered by low pay, geographic and social isolation, and challenging work conditions (Hammer et al., 2005). An undersupply of highly effective teachers would limit the strategies available to rural schools and districts for improving their teacher workforce through staffing practices.

CONTACT Erica Harbatkin harbatki@msu.edu Michigan State University, Education Policy Innovation Collaborative, 236 Erickson Hall | 620 Farm Lane, East Lansing, MI 48824, United States

Under the Every Student Succeeds Act (ESSA), all states are required to identify their lowest performing schools for Comprehensive Support and Improvement (CSI). Unlike school turnaround under RTTT and SIG, ESSA calls for states to tailor interventions in their lowest performing schools to local needs and does not require staff replacement strategies. Still, teacher quality plays an outsized role in student achievement, and in particular a stable and effective teacher workforce is critical to student performance on accountability exams – which contribute heavily to whether a school is classified as low performing (see, e.g., Aaronson et al., 2007; Chetty et al., 2014; Rockoff, 2004). To that end, states and districts under ESSA are focused on improving teacher quality, but they are likely taking a variety of approaches to doing so. Understanding whether rural schools have the resources to dismiss or coach out their least effective teachers and whether they are able to recruit more effective teachers than they lose would provide needed context for improvement planning in rural turnaround schools.

In this paper, I try to understand differences between staffing for low-performing schools in rural compared with non-rural settings. These differences are important for educators and policymakers in states with populations of struggling rural schools to understand as they work to answer the call to address staffing challenges unique to rural low-performing schools (Ayers, 2011; Player, 2015; Stoddard & Toma, 2021). Specifically, this study examines the case of staffing for turnaround in rural schools through a descriptive analysis of the first cohort of Comprehensive Support and Improvement (CSI) schools in North Carolina, where about half of schools are located in rural settings. I ask two questions.

(1) To what extent do teachers in rural CSI schools differ from teachers in non-rural CSI schools at the time of identification?
(2) To what extent does teacher hiring, assignment, and retention appear to vary between rural and non-rural CSI schools.

The goal of this analysis is to provide a snapshot in time of teacher labor markets in rural schools and the ways these schools are approaching staffing given their labor markets. It is not intended as a causal analysis and does not speak to whether CSI identification or implementation induced schools to make particular staffing decisions. Rather, it is intended to build up the nascent literature on staffing in rural low-performing schools.

Literature Review

Given the importance of effective teachers for student learning, some of the most pervasive federal education policies of the 21st century have centered around improving teacher quality, from teacher evaluation reforms to school turnaround interventions aimed at improving the quality of the workforce in low-performing schools under the Elementary and Secondary Education Act (Bleiberg & Harbatkin, 2020; Close et al., 2018; Howell & Magazinnik, 2017). However, there is reason to believe that improving educator quality may require different approaches depending on school contextual factors, including school performance, poverty, and urbanicity. While an array of studies address staffing in the context of low school performance and high poverty, we know little about the role of urbanicity in staffing for low-performing schools. In this literature review, I highlight this gap through an overview of the current research on educator quality in low-performing schools including accountability-driven staffing strategies, the research on rural teacher labor markets, and finally the limited research on teacher labor markets in rural low-performing schools.

Policies aimed at improving teacher quality in low-performing schools have shown promise toward successful school improvement. Under RTTT and SIG, turnaround interventions focused on improving educator quality through staff replacement were more effective than other interventions (Carlson & Lavertu, 2018; Dee, 2012; Strunk et al., 2016; Sun et al., 2017). Studies showing positive effects of school turnaround also found increased retention of effective teachers – while studies identifying negative effects found the intervention increased teacher turnover (Heissel & Ladd, 2018; Henry &

Harbatkin, 2020; Papay & Hannon, 2018; Sun et al., 2017). One study found turnover of effective teachers explicitly suppressed the effects of a turnaround intervention in Tennessee, (Henry et al., 2020). This research is largely focused on initiatives dominated by urban settings, such as District of Columbia Public Schools (DCPS), Memphis, Houston, Los Angeles, and San Francisco (Adnot et al., 2017; Henry et al., 2020; James & Wyckoff, 2020; Strunk et al., 2016; Sun et al., 2017).

In high-stakes accountability settings, leaders of low-performing schools and districts have employed a variety of strategies to improve teacher quality. Some of these strategies have involved efforts to recruit and retain highly effective teachers, for example, through hiring and performance bonuses (Balch & Springer, 2015; Pham et al., 2021). Others have aimed to dismiss teachers deemed less effective, for example, through tenure reform, or to develop less effective teachers through professional development or coaching (Allen et al., 2011; Bastian & Marks, 2017; Brunner et al., 2019; Early et al., 2017; Goldhaber & Hansen, 2010; Loeb et al., 2015; Papay & Hannon, 2018). Policies aimed at dismissing less effective teachers typically aim to hire more effective teachers to replace them with the goal of creating a positive compositional effect of the turnover (Adnot et al., 2017; Cullen et al., 2019; Dee & Wyckoff, 2015). When these efforts are successful, they can elevate teaching quality across the school (see, e.g., Early et al., 2017; Papay & Hannon, 2018). In other instances, education leaders have taken more targeted approaches to staffing by aiming to increase teacher quality specifically in grades and subjects subject to accountability. In these cases, schools have concentrated their most effective teachers in tested grades and subjects while assigning their less effective teachers to untested grades or subjects rather than developing or dismissing them (Chingos & West, 2011; Cohen-Vogel, 2011; Fuller & Ladd, 2013; Goldring et al., 2015; Grissom et al., 2017; Kraft et al., 2020). For schools with less robust educator labor markets or fewer resources to develop new teachers, these targeted staffing strategies may be particularly appealing to avoid a low-performing designation or make quick, short-term improvements to student achievement in grades and subjects with test scores that determine consequences. One context in which this might be true is rural schools. Additionally, these targeted staffing strategies may be especially appealing in cases of high unintentional turnover. When schools are already struggling with high rates of turnover, they may not have the capacity to dismiss additional teachers and create even more teacher vacancies (Rice & Malen, 2003).

There is a growing research base focused broadly on teacher turnover and rural labor markets, highlighting important differences between rural and non-rural – including urban and suburban – teacher labor markets (see. e.g., Ayers, 2011; Hammer et al., 2005; Player, 2015; Stoddard & Toma, 2021). Rural schools face particular challenges recruiting and retaining teachers due to a limited labor supply, geographic isolation, lower pay, and fewer staff positions leading to more responsibilities for each teacher (Barrow & Burchett, 2001; Berry & Gravelle, 2013; Brownell et al., 2005; Hammer et al., 2005, p. 2005; Miller, 2012; Player, 2015; Rosenberg et al., 2015). A study of staffing in California schools showed rural schools have more vacancies and rely more on emergency credentialed teachers than schools located in urban, suburban, and town locales (Goldhaber et al., 2020). The shallow labor pool likely stems at least in part from the fact that teachers are more likely to transfer out of rural schools than into them, and that teachers in rural schools are more likely to exit the profession entirely than their non-rural counterparts (Cowen et al., 2012). These patterns suggest rural schools have a harder time retaining effective teachers and recruiting effective replacements to fill vacancies. Given the challenges recruiting and retaining teachers in low-performing and high poverty schools in particular, it stands to reason that rural low-performing schools face unique staffing challenges due to their context.

However, very little is known about staffing for rural low-performing schools, though there is evidence that recruiting and retaining a sufficient number of highly effective teachers may be especially challenging in this context (Rosenberg et al., 2015). As a result, school turnaround interventions aiming to improve the quality of the teacher workforce through staff replacement may be unrealistic in the rural context. There is evidence that these school turnaround strategies focused on reconstituting the teaching staff are not feasible in schools with limited resources and local labor markets (Malen et al., 2002; Malen & Rice, 2016; Rice & Malen, 2003). Indeed, one study of the effect of a largely rural turnaround intervention found that less effective teachers were assigned at a higher rate to untested early grades (Henry et al., 2021).

Theoretical Framework

According to economic labor market theory, the demand for teachers can be characterized as the number of teaching positions in a given context and the supply is the available pool of teachers willing to work in those available positions (Guarino et al., 2006). Two relevant school contextual factors affecting the supply of teachers are school urbanicity and the socioeconomic background and achievement of the student body. In particular, rural areas tend to have fewer qualified teachers in their local labor markets than urban and suburban settings (Hammer et al., 2005; Jimerson, 2005; McClure & Reeves, 2004), and schools with high rates of students in poverty and low-performing students tend to have fewer applicants than more affluent, higher performing schools (McClure & Reeves, 2004).

As a result, the supply of qualified teachers may not be sufficient given the demand in rural and low-performing schools. While teacher shortages are not universal (Ingersoll, 2001, 2003) several studies have documented severe shortages in rural and low-performing schools (e.g., Boyd et al., 2005a; Loeb & Reininger, 2004). These findings suggest that although states may have a large supply of certified teachers, teachers may have preferences diverting them away from positions in certain schools or they may lack relevant information about all positions (Boyd et al., 2005b; Cannata, 2010; Reininger, 2012). Though there is little research on low-performing rural schools, related research above provides cause for concern that staffing shortages would be especially stark in this subset of schools, which face challenges related to both locale and school socioeconomics.

There are several explanations for why the supply of teachers in rural schools may be insufficient to meet demand. These include school-specific factors such as funding formulas resulting in lower pay combined with less favorable working conditions including more responsibilities and more preps (Maranto, 2013; Shuls, 2018). Factors outside of the school include fewer highly qualified teachers living nearby as well as fewer teachers interested in moving to rural areas, where they may experience geographic and social isolation and which can lack the amenities of urban and suburban areas (Collins, 1999; Proffit et al., 2004). In low-performing schools in general, the supply of teachers may be insufficient due to challenging working conditions, lower pay, and concerns and uncertainty related to high-stakes accountability (Boyd et al., 2005a; Clotfelter et al., 2004).

To the extent that demand outpaces supply, schools will struggle to fill open vacancies and may have to fall back on less optimal approaches to filling these vacancies, such as emergency certification, long-term substitutes, larger class sizes, and fewer course offerings (Brownell et al., 2005; Jimerson, 2005; Mobra & Hamlin, 2020). If supply just barely meets demand, there will be little competition for these positions and schools will have fewer choices to fill available slots. As a result, schools with limited teacher supply will not have the bandwidth or resources to compete for the most effective teachers. Thus, schools may need to replace teachers who turn over with teachers who are similarly or even less effective. Because teacher turnover is associated with decreased student achievement in the short term (Ronfeldt et al., 2013), even replacing a departing teacher with one who is exactly as effective could have negative effects on student achievement.

A central tenet of school turnaround is improving educator quality. I consider two ways low-performing schools may achieve this goal. The first is replacing teachers who leave with more effective replacements. Under RTTT and SIG, this strategy was central to the federally defined turnaround model, which called for low-performing turnaround schools to replace at least 50% of their staff. However, if rural schools do not have robust labor markets from which to recruit highly effective teachers, they may end up replacing ineffective teachers with other ineffective teachers – creating teacher churn that actually decreases student achievement due to the negative disruption effect of the turnover. The second approach I consider is a strategic staffing practice in which schools reassign ineffective teachers away from tested grades and subjects while retaining effective teachers in (or reassigning them to) tested grades and subjects (Chingos & West, 2011; Cohen-Vogel, 2011; Fuller & Ladd, 2013; Goldring et al., 2015; Grissom et al., 2017; Kraft et al., 2020). This latter approach may yield quick upticks in accountability exam scores while undermining learning in untested grades and subjects and potentially hindering the sustainability of any turnaround. However, if rural low-performing schools lack an adequate supply of effective teachers to fill vacancies, the high-stakes accountability setting may induce school leaders to engage in these practices to avoid sanctions or being labeled as low performing.

School Turnaround in North Carolina

North Carolina has a rich history of school turnaround policy, beginning with a court-ordered turnaround in 128 low-performing schools from 2006 through 2010 (Thompson et al., 2011). The state's RTTT intervention followed in 118 schools from 2010 through 2015, and produced mixed effects on student achievement (Heissel & Ladd, 2018; Henry & Guthrie, 2019; Henry et al., 2015). When RTTT funds ran out, the state continued its work in a smaller group of low-performing schools as part of a new turnaround intervention, the North Carolina Transformation initiative (NCT) from 2015–16 to 2016–17. By excluding the 10 largest school districts in the state, NCT focused almost exclusively on rural schools. These largely rural NCT schools experienced higher teacher turnover than similar low-performing schools, but the turnover did not improve the quality of the NCT teacher workforce. In fact, student achievement decreased after two years of turnaround under NCT (Henry & Harbatkin, 2020).

The state identified its CSI schools in October 2018 using 2017–18 school performance data. Schools were targeted for comprehensive support and improvement if they scored in the bottom 5% of all Title 1 schools on the state accountability index (which includes proficiency rates, value-added, English Learner progress, and high school graduation rates) or were a high school with four-year cohort graduation rate lower than 66.7%. The state's comprehensive school reform model calls for districts to carry out needs assessments in these schools and to tailor reform strategies to the unique needs of the school. Districts and schools then develop improvement plans aligned to their local needs and identify strategies for meeting targets set in those plans. A report to the state Board of Education identified high teacher turnover and need for teacher development as a common theme in improvement plans, underscoring that low-performing schools and districts appreciated the pertinence of teacher quality to school improvement (North Carolina State Board of Education, Department of Public Instruction, 2020). Teacher replacement is not a required component of school reform under North Carolina's ESSA model, though districts may choose to implement a RTTT/SIG reform model in their low-performing schools, which would involve teacher replacement.

Data and Methods

Sample and Data

This analysis draws from a longitudinal database of statewide administrative data maintained by the University of North Carolina-Chapel Hill's Educational Policy Initiative at Carolina (EPIC). The database contains data on all North Carolina public schools. I restrict the sample to the 94 traditional public schools[1] on the 2018–19 CSI list, which includes 43 rural and 51 non-rural schools. This list includes schools identified for low performance and for low graduation rates as described in the section above.[2] The analytic sample includes teachers in schools designated for low performance or low graduation rates. I draw from three years of data from 2017–18 (the year that schools were in the bottom 5% or had low graduation rates), 2018–19 (the identification year), and 2019–20 (the first year of full implementation). Specifically, I use teacher effectiveness data from 2017–18 and 2018–19 to examine staffing changes in rural and non-rural schools in 2018–19 and 2019–20. Because the state suspended testing in response to the COVID-19 pandemic, 2019–20 data are not fully available (there are no state tests and I am unable to identify which teachers are teachers of tested subjects and grades; I therefore use 2018–19 teacher data to examine teacher assignments.). The outcome analyses include 4,905 teacher-year-observations in 2018–19 and 2019–20 representing 3,270 unique teachers. I classify schools as rural or non-rural using U.S. Census locale codes.[3]

Measures

Broadly, strategic staffing practices comprise two related components – (a) teacher quality and experience, and (b) teacher mobility and assignments. The remainder of this section describes the measures I use for each of those components.

Teacher Quality and Experience

I draw from two types of teacher quality measures to understand how rural and non-rural schools leverage available human capital to improve instructional quality. The first is teacher value-added scores measured by the Education Value-Added Assessment System (EVAAS), which provide effectiveness measures for teachers of tested grades and subjects as well as ELA teachers in kindergarten through third grade (the latter is not a tested grade and subject; instead these scores are generated from formative assessment data). The EVAAS measure for untested teachers therefore draws from performance on benchmark assessments. The state calculates EVAAS value-added from student assessment scores as a function of the students' prior scores in the tested subject. EVAAS scores are a continuous measure that can theoretically range from negative to positive infinity (Wright et al., 2010). In other research, value-added scores using these estimation methods have been shown to be valid measures of effectiveness that identify teachers who produce higher achievement among their students (Kane et al., 2013). Because I only draw EVAAS scores from tested grades and subjects, as well as K-3 reading comprehension, they are available for slightly more than one-third of the sample in 2018 and 2019. I examine staffing decisions based on *lagged* EVAAS scores; in other words, I am interested in 2017–18 scores for 2018–19 turnovers. I observe lagged EVAAS scores for 30% of the sample.

I generate three different effectiveness measures using EVAAS. The first is a teacher-level mean for all subjects and classes the teacher taught in a given year. In the full sample of North Carolina schools during the study period, the mean of this teacher-level mean is 0.09 in each year with a standard deviation of 2.3. EVAAS scores are lower in CSI schools than in higher performing schools. The mean teacher-level EVAAS score in these schools across the two years is -1.05, with a standard deviation of 2.3. The second value-added measure is an indicator denoting whether the teacher is highly effective. Specifically, teachers receive one of three ratings based on their EVAAS score for a given subject: they *meet expected growth* if they are within two points of predicted growth on the EVAAS scale, *exceed expected growth* at more than two points above, and *do not meet expected growth* at more than two points below. Following these state definitions, I code a teacher as highly effective if they exceed expected growth in any subject or class. I then create a third measure to classify ineffective teachers as those who do not meet expected growth in one or more subjects. These value-added scores capture just one dimension of teacher effectiveness and are specific to student test scores. However, they are relevant in the context of staffing because the state makes them available to principals. To that end, they are one dimension of teacher effectiveness principals can draw from in making staffing decisions.

Next, I draw a set of teacher effectiveness measures from the teacher's annual evaluation measured by the North Carolina Educator Effectiveness System (NCEES), which includes five standards: (1) teacher leadership, (2) establishing a respectful learning environment for diverse students, (3) content knowledge, (4) facilitation of learning for students, and (5) reflection on practice. These scores draw from observations of teacher practice and theoretically provide information on dimensions of teacher effectiveness beyond the information in EVAAS scores. Teachers receive scores of 1–5 on each rating, with one being the lowest rating a teacher can receive and five the highest. Teachers with up to three years of experience are evaluated on all standards, while more experienced teachers are only required to be evaluated on standards 1 and 4 (though their districts can choose to evaluate them on all 5). Full NCEES scores (i.e., scores on all 5 dimensions) are available for approximately 60% of the sample. Partial NCEES scores (i.e., scores on standards 1 and 4) are available for about 85% of the sample. I draw three measures from the teacher NCEES score – the median score for just those teachers evaluated on all five standards, the score on standard 1, and the score on standard 4. Again, I examine staffing decisions based on lagged NCEES scores. I can observe a lagged median score for 47% of the sample and lagged scores on standards 1 and 4 for 69% of the sample. The NCEES measures have the benefit of covering more of the sample than the EVAAS measures, but have substantially less variation (more than 90% of the sample receives a 3 or 4) and are more subjective.

Finally, I create two variables related to teacher experience. The first is the number of years of teaching experience, ranging from zero for new teachers to a maximum of 48 in the sample, with a mean of approximately 10 and a median of eight. The second is an indicator denoting whether the teacher has three or fewer years of experience and is therefore considered "novice" in the state. About one-third of teachers in the sample are novice and two-thirds are experienced.

Teacher Mobility and Assignments

Decisions around teacher mobility and assignments include hiring, retention, dismissal (or coaching out), and assigning teachers to classes. Ideally, schools want to hire and retain the most effective teachers, and dismiss the least effective teachers. In the context of school-level turnaround, the relevant variables are therefore teacher turnover from the school and new-to-school teachers. I code a teacher as a turnover if they leave the school during or at the end of the school year. I code a teacher as new-to-school if they are new to the school in a given school year. Just under one-third of teachers in the sample turn over and are new to school. I focus on teacher turnover in 2018–19 and new-to-school teachers in 2019–20.

I identify tested teachers as those who have students taking an exam counting toward the school's state accountability score. In North Carolina, these exams are third- through eighth-grade math and reading end-of-grade (EOG) exams, fifth and eighth grade science EOGs, and end-of-course (EOC) exams in high school grades – specifically Math 1, English 2, and Biology. About 18% of teachers in the sample teach in a tested course. This figure is slightly higher in rural schools than non-rural schools – potentially because rural schools have fewer resources to offer courses outside of the academic core. Ideally, I would draw the tested teacher variable from 2019–20, but I cannot create such a measure due to the lack of accountability exams in 2020. I instead create the tested teacher variable using 2018–19 data. It is therefore important to consider the results related to tested subjects in light of the timing of the measure. At the time when districts and schools were making teacher assignments, they would have known the school was relatively low performing based on student test scores and graduation rates, but would not have yet received their CSI classification. As such, the tested teacher variable reflects a response to accountability pressures, but not specifically to the CSI designation as the turnover and new-to-school variables do.

Empirical Strategy

I begin with a descriptive analysis of the effectiveness and experience levels of teachers in rural and non-rural turnaround schools in North Carolina. Specifically, I use *t*-tests to compare characteristics of teachers in rural compared with non-rural turnaround schools. This analysis provides context on the educator labor markets in these settings pooled over three years from 2017–18 through 2019–20. I then examine whether rural and non-rural schools appeared to employ different teacher staffing and assignment strategies by testing the effectiveness and experience levels of teachers who (a) turn over, (b) are new to school, and (c) are assigned to tested grades and subjects in rural relative to non-rural schools. To do so, I run a series of descriptive ordinary least squares (OLS) regressions predicting teacher effectiveness or experience as a function of teacher assignment and the school's urbanicity.[4]

I run these models to predict the outcomes related to lagged teacher value added (EVAAS, exceeded growth, did not meet growth), teacher evaluation (median NCEES, teacher leadership NCEES standard, student learning NCEES standard), and teacher experience (years of teaching experience, novice teacher). Specifically, I run three OLS regressions, with the models predicting EVAAS taking the form

$$EVAAS_{ist-1} = \beta_0 + \beta_1 Rural_s + \beta_2 Turnover_{it} + \beta_3 Rural_s \times Turnover_{it} + \sigma + \varepsilon_{ijst} \tag{1}$$

$$EVAAS_{ist-1} = \beta_0 + \beta_1 Rural_s + \beta_2 NewToSchool_{it} + \beta_3 Rural_s \times NewToSchool_{it} + \sigma + \varepsilon_{ist} \tag{2}$$

$$EVAAS_{ist-1} = \beta_0 + \beta_1 Rural_s + \beta_2 Tested_{it} + \beta_3 Rural_s \times Tested_{it} + \sigma + \varepsilon_{ist} \tag{3}$$

where Equation 1 predicts the EVAAS score for teacher *i* in school *s* in year *t*-1 as a function of whether the school is rural, whether the teacher turns over from the school, an interaction between the rural and turnover indicators, a school level fixed effect denoted as σ (elementary, middle, high with elementary as the reference category) and an idiosyncratic error term clustered at the school level. A negative estimate on β_2 would provide evidence that schools turned over less effective teachers than they retained. Differences by school urbanicity would be apparent in β_3; a positive estimate on β_3 would indicate that rural turnaround schools lost more effective teachers than non-rural turnaround schools.

In Equation 2, I replace the turnover indicator with a new-to-school indicator to examine the extent to which these turnaround schools recruited more effective (or experienced) teachers than they had on staff. Here, a negative estimate on β_3 would suggest that rural turnaround schools recruited less effective teachers than non-rural turnaround schools. Finally, in Equation 3, I replace the indicator of interest with a variable indicating whether a teacher is in a tested grade and subject. To the extent that rural schools engaged in more frequent strategic staffing in the form of assigning effective teachers to tested grades and subjects, it would be evident in β_3.

The models for teacher evaluation (NCEES) and teacher experience follow a parallel form, though I draw the teacher experience measure from the same year as the teacher assignment rather than the lagged year. The models predicting a dichotomous measure (i.e., exceeds growth, does not meet growth, novice teacher) are linear probability models and can therefore be interpreted in percentage points. In other words, in Equation 1 predicting "exceeds growth," β_2 provides the difference in probability that a teacher who turned over exceeded growth relative to a teacher who did not turn over, and β_3 provides the second difference for a teacher in rural schools relative to non-rural schools.

Finally, I rerun the NCEES and experience models separately by school level (elementary, middle, and high, dropping the school level fixed effects). I do not run these school-level models using EVAAS measures due to small sample sizes in high school in particular, where very few teachers have lagged EVAAS scores. I present these tables in an appendix and highlight meaningful differences at the end of the Results section.

While the lagged teacher effectiveness measure on the left side of the equation may seem counter-intuitive, setting up the regressions in this way allows for significance testing of observed descriptive patterns by subgroups. The results are descriptive and should be interpreted as associational and not as causal effects. It is also important to note that these results represent a snapshot in the first year of CSI implementation. They do not provide evidence of longer term trends in these schools.

Results

I begin this section with a description of teacher effectiveness and experience in rural and non-rural turnaround schools to characterize the differences in teacher labor markets across contexts. I then move to the regression results showing differential staffing practices by school locale. I conclude with a brief summary of results by school level.

Teacher Effectiveness and Experience by Setting

Table 1 shows the results of -tests comparing teachers in rural and non-rural CSI schools in the analytic sample. Row 1 shows that the mean teacher-level value-added score was -1.36 in rural schools relative to -0.81 in non-rural schools – a difference of more than 0.2 standard deviations. Rows 2 and 3 show that teachers in rural turnaround schools were 2.5 percentage points less likely t-t to exceed growth and nearly 7 percentage points more likely not to meet growth than teachers in non-rural turnaround schools. These patterns were reversed in the NCEES teacher evaluation measures, with teachers in rural turnaround schools scoring slightly higher on all three evaluation measures than their counterparts in non-rural turnaround schools – though the differences on these measures are small and may be driven by site-level differences in evaluators rather than true differences in teacher quality. In terms of teacher experience, teachers in rural settings had about half a year more experience, on average, than teachers in non-rural settings, though there were a similar number of novice and experienced teachers across the two locale types. The final panel shows that teachers in rural schools were similar in age and sex to teachers in non-rural schools, but more likely to be White and less likely to be Black or Hispanic than their peers in non-rural schools. In sum, this analysis shows that schools in rural settings appeared to have less effective teachers in tested subjects and a less diverse teacher workforce than schools in non-rural settings, while rural and non-rural schools are relatively similar on other measures.

Staffing for School Improvement

Tables 2 through 4 show the estimates for the descriptive regressions predicting the value-added, NCEES teacher evaluation, and teacher experience measures, respectively. The columns identify the outcome variable (i.e., for Table 2, teacher value added, dichotomous indicator for exceeds growth, and dichotomous indicator for does not meet growth). The different panels are for the teacher mobility or assignment indicator on the right side of the equation (i.e., turnover in Panel A, new to school in Panel B, and tested subject in Panel C). In each panel, the first row provides the coefficient estimate on the rural school indicator, which reflects the estimated value-added score, probability of

Table 1. Teacher effectiveness, experience, and teaching assignment by school urbanicity.

	Rural	Non-rural	Difference	N
Teacher effectiveness				
EVAAS teacher mean score	−1.356	−0.806	−0.550***	1,827
Exceeds growth	0.065	0.090	-0.025^{+}	1,827
Does not meet growth	0.363	0.291	0.072**	1,827
NCEES standard 1	3.476	3.391	0.086***	4,170
NCEES standard 4	3.390	3.273	0.117***	4,170
NCEES median (full evals only)	3.291	3.194	0.097***	2,909
Teacher experience				
Teaching experience	10.657	10.039	0.617**	7,320
Novice teacher	0.328	0.331	−0.003	7,320
Experienced teacher	0.672	0.669	0.003	7,320
Teacher mobility and assignment				
Teacher turnover	0.310	0.315	−0.005	4,932
New to school	0.298	0.292	0.006	4,905
Teaches tested subject	0.201	0.170	0.031***	7,397
Teacher demographics				
Female	0.760	0.758	0.002	4,838
Age	40.764	39.869	0.895*	4,858
White	0.510	0.426	0.084***	4,613
Black	0.432	0.511	−0.078***	4,613
Hispanic	0.025	0.044	−0.019***	4,613
Asian	0.009	0.011	−0.002	4,613
Other	0.023	0.007	0.016***	4,613

Estimates from *t*-tests. $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 2. Teacher value-added (EVAAS).

	(1) Value-added	(2) Exceeds growth	(3) Does not meet
Panel A – Estimated value-added for teachers who turn over (2018-19)			
Rural	-0.615*	-0.047^{+}	0.082
	(0.300)	(0.024)	(0.057)
Teacher turnover	-0.542*	-0.028	0.131**
	(0.265)	(0.030)	(0.047)
Rural X Teacher turnover	0.745^{+}	0.040	-0.161^{+}
	(0.437)	(0.037)	(0.086)
Observations	*733*	*733*	*733*
Panel B – Estimated value-added scores for new-to-school teachers (2019-20)			
Rural	-0.630*	-0.021	0.091^{+}
	(0.315)	(0.031)	(0.054)
New to school	0.285	0.022	-0.067
	(0.397)	(0.043)	(0.055)
Rural X New to school	0.637	0.109	-0.020
	(0.655)	(0.095)	(0.092)
Observations	*752*	*752*	*752*
Panel C – Estimated value-added scores for tested teachers (2018-19)			
Rural	0.060	-0.003	0.013
	(0.407)	(0.045)	(0.071)
Teaches tested subject	0.139	-0.001	0.018
	(0.207)	(0.036)	(0.041)
Rural X Teaches tested subject	-0.724	-0.053	0.020
	(0.472)	(0.055)	(0.084)
Observations	*733*	*733*	*733*

Note: Standard errors clustered at the school level. All models include school type fixed effects. Turnover measured in 2018-19, new-to-school measured in 2019-20, and tested classification based on in 2018-19 teaching assignment. EVAAS measured in the year prior to the relevant pathway (i.e., 2017-18 in Panels A and C and 2018-19 in Panel B). $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

exceeding growth, and probability of not meeting growth, respectively, for teachers in rural schools. The second row provides the coefficient estimate on the teacher mobility or assignment indicator of interest for a given panel. The third row provides the estimated difference in effectiveness between rural and non-rural schools. Across all tables, the primary estimate of interest is the coefficient in Row 3 because it shows the difference in effectiveness (Tables 2 and 3) or experience (Table 4) between teachers in rural and non-rural schools, while the coefficient in Row 2 provides the estimate for non-rural schools.

Row 1, Column 1, shows that teachers in rural schools had value-added scores that were 0.62 EVAAS units lower – conditional on school type fixed effects – than teachers in non-rural schools. Row 2 shows that in non-rural schools (the reference category), teachers who turned over had value-added scores that were 0.54 EVAAS units lower than teachers who stay, again conditional on school type fixed effects. In other words, in non-rural turnaround schools, teachers who left had value-added scores that were about 0.23 standard deviations lower than teachers who stayed. However, Row 3 shows that teachers who turned over in rural schools were 0.75 EVAAS units – about one-third of a standard deviation – *more effective* than teachers who stayed, relative to non-rural schools. This estimate indicates that rural schools lost more effective teachers than non-rural schools. Put another way, the teachers who left rural schools were *more* effective than the teachers who stayed, while the teachers who left non-rural schools were *less* effective. Column 2, Row 1 shows that teachers in rural schools were less likely to have previously exceeded growth, while the insignificant estimate in row 3 shows rural schools were not significantly more likely than non-rural schools to lose their highly effective teachers. Descriptively, however, the teachers who turned over were more likely to be highly effective in rural than non-rural schools, showing that rural schools often did not retain the small

Table 3. Teacher evaluation (NCEES).

	(1) Standard 1	(2) Standard 4	(3) NCEES median
Panel A – Estimated evaluation scores for teachers who turn over (2018-19)			
Rural	0.173**	0.175**	0.194**
	(0.064)	(0.066)	(0.072)
Teacher turnover	-0.172***	-0.158***	-0.133**
	(0.042)	(0.046)	(0.041)
Rural X Teacher turnover	0.053	0.075	0.063
	(0.075)	(0.082)	(0.081)
Observations	1,713	1,713	1,176
Panel B – Estimated evaluation scores for new-to-school teachers (2019-20)			
Rural	0.024	0.069	0.009
	(0.064)	(0.065)	(0.066)
New to school	0.036	0.146*	-0.012
	(0.075)	(0.061)	(0.073)
Rural X New to school	-0.238+	-0.297*	-0.162
	(0.124)	(0.118)	(0.101)
Observations	1,710	1,710	1,147
Panel C – Estimated evaluation scores for tested teachers (2018-19)			
Rural	0.153*	0.166*	0.202*
	(0.070)	(0.071)	(0.079)
Teaches tested subject	-0.041	-0.041	-0.007
	(0.049)	(0.049)	(0.052)
Rural X Teaches tested subject	0.131+	0.116	0.052
	(0.076)	(0.079)	(0.085)
Observations	1,713	1,713	1,176

Note: Standard errors clustered at the school level. All models include school type fixed effects. Turnover measured in 2018-19, new-to-school measured in 2019-20, and tested classification based on in 2018-19 teaching assignment. NCEES is measured in the year prior to the relevant pathway (i.e., 2017-18 for Panels A and C, and 2018-19 for Panel B). Most teachers receive scores for Standard 1 and 4 as part of the abbreviated assessment. Teachers with three or fewer years of experience and some other teachers, such as those up for license renewal, receive scores for all five standards as part of the comprehensive assessment. The models predicting median rating restrict the sample to just those teachers with scores on all five standards. $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

number of highly effective teachers that they were able to recruit. Finally, Column 3 shows that while teachers who left non-rural schools were 13.1 percentage points *more likely* to be ineffective – their counterparts who left rural schools were 16.1 percentage points *less likely* than they were to be ineffective, though the latter coefficient is only marginally significant at the $p < .10$ level. Again, this significant difference in effectiveness between teachers who left rural and non-rural schools highlights a difference in staffing practices by school locale in the observed year. While I find evidence suggesting non-rural turnaround schools may have used teacher value-added data to make teacher dismissal decisions, I do not find that rural turnaround schools employed the same strategy.

I do not find evidence that either rural or non-rural turnaround schools hired more effective teachers than they already had on staff (Panel B) or that either set of schools placed more effective teachers in tested subjects (Panel C). These null findings should be interpreted in light of the highly imprecise estimates for both sets of models. The Panel B estimates are imprecise, in part at least, because new-to-school teachers are more likely to be first-year teachers and less likely to have lagged value-added scores. For Panel C, it is useful to recall that the teacher value-added scores largely come from tested subjects. Thus, few teachers not already in tested subjects would have a value-added score for the prior year. I therefore turn next to teacher effectiveness definitions based on the state's comprehensive teacher evaluation system, which provides evaluation scores for a larger portion of the teacher workforce.

Table 3 provides the results for the same set of outcome measures (i.e., teacher turnover, new-to-school teacher, and tested teacher) using teacher effectiveness measures from the comprehensive teacher evaluation system, NCEES. Column 1 provides the estimate for Standard 1 measuring

Table 4. Teacher experience.

	(1) Years of experience	(2) Novice teacher
Panel A – Estimated experience for teachers who turn over (2018-19)		
Rural	0.173**	0.175**
	(0.064)	(0.066)
Teacher turnover	-0.172***	-0.158***
	(0.042)	(0.046)
Rural X Teacher turnover	0.053	0.075
	(0.075)	(0.082)
Observations	1,713	1,713
Panel B – Estimated experience for new-to-school teachers (2019-20)		
Rural	0.024	0.069
	(0.064)	(0.065)
New to school	0.036	0.146*
	(0.075)	(0.061)
Rural X New to school	-0.238^{+}	-0.297*
	(0.124)	(0.118)
Observations	1,710	1,710
Panel C – Estimated experience for tested teachers (2018-19)		
Rural	0.153*	0.166*
	(0.070)	(0.071)
Teaches tested subject	-0.041	-0.041
	(0.049)	(0.049)
Rural X Teaches tested subject	0.131^{+}	0.116
	(0.076)	(0.079)
Observations	1,713	1,713

Note: Standard errors clustered at the school level. All models include school type fixed effects. Turnover measured in 2018-19, new-to-school measured in 2019-20, tested classification measured in 2018-19. Novice teachers are those with up to 3 years of teaching experience. Experienced teachers are those with more than 3 years of teaching experience. $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

teacher leadership, Column 2 provides the estimate for Standard 4 measuring student learning, and Column 3 provides the median score across all five standards for just those teachers with full assessment scores. Panel A shows that the teachers who turned over were less effective across all three measures in both non-rural and rural schools, suggesting the teachers who turned over in both locale types were less effective than the teachers who stayed. Specifically, Column 1, Row 2 shows that in non-rural schools, the teachers who turned over had a score of 0.17 points lower on the five-point NCEES scale than the teachers who stayed, and the insignificant estimate in Row 3 shows that teachers who left rural schools were not significantly more or less effective than teachers who left non-rural schools. However, the third row of Panel B shows that the new-to-school teachers in rural schools were less effective than the new-to-school teachers in non-rural schools, relative to the teachers who stayed. Taken together, Panels A and B show that while rural schools turned over less effective teachers than they retained, the new-to-school teachers appear to be even less effective than the teachers they replaced.

Finally, Panel C provides descriptive evidence that rural schools assigned more effective teachers (as measured by evaluation scores) than non-rural schools to a tested grade and subject. Specifically, Row 3 shows that the Standard 1 score (teacher leadership) was 0.13 points higher on the NCEES scale (about one-fifth of a standard deviation) for tested teachers than untested teachers in rural schools relative to non-rural schools. This pattern is similar across all three teacher evaluation measures but only marginally significant for Standard 1. In other words, teachers of tested subjects in rural schools received comparably higher evaluation ratings than teachers of untested subjects – while teachers of tested subjects in non-rural schools received similar evaluation ratings to their untested peers. The lack of statistical significance across all three measures could indicate that the consistent pattern appears only by chance, though it could also stem from the very small variation in the NCEES variable; indeed, the measure with the most variation is standard 1, where the most precise estimates here emerge. This finding provides some descriptive evidence suggesting rural schools leveraged the talent they had and were able to retain in order to focus on tested grades and subjects.

Figure 1 brings together the findings from the two measures of teacher effectiveness by showing the estimated lagged value-added score for each teacher pathway, with pathways related to mobility in the left panel and pathways related to classroom assignment in the right panel. Markers represent marginal predictions of the lagged teacher effectiveness score for teachers in rural and non-rural

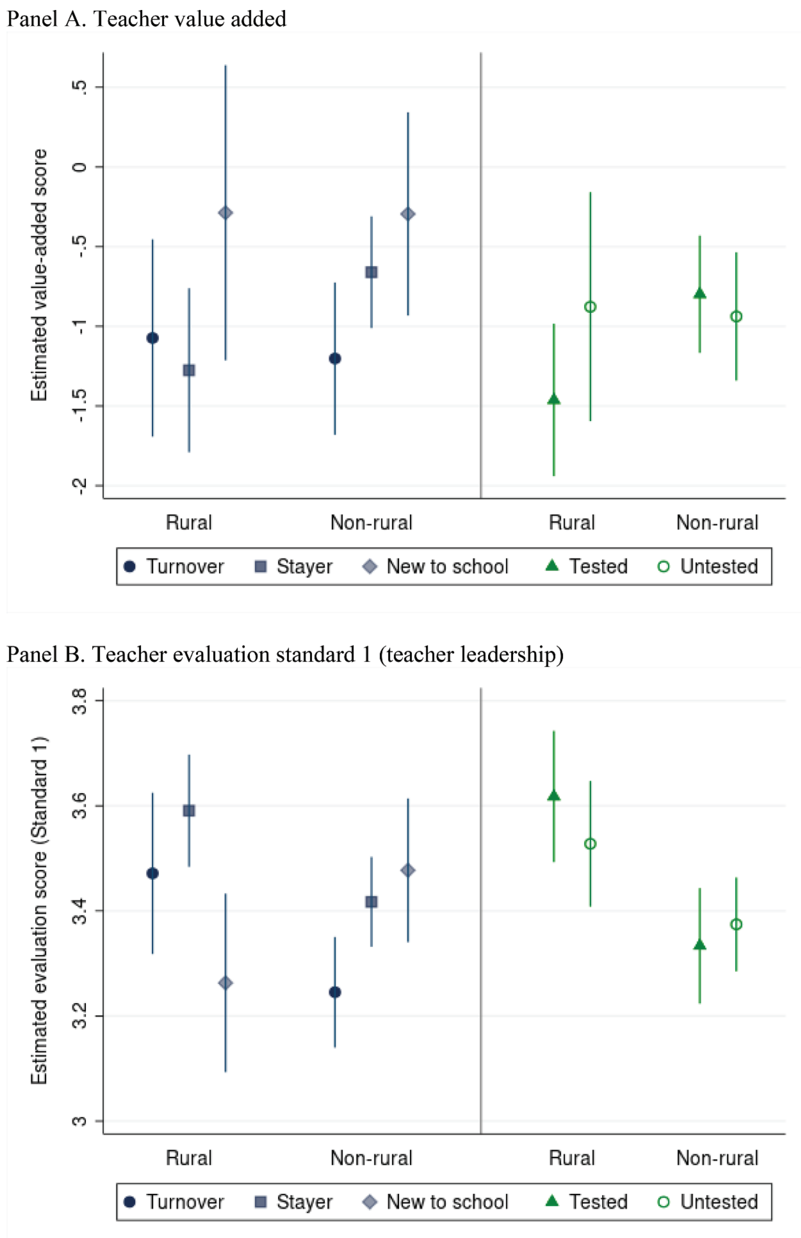

Figure 1. Marginal estimates of teacher effectiveness scores by mobility and assignment in following year. Markers represent marginal effect estimates for value-added score (Panel A) and evaluation score on NCEES standard 1 (Panel B). Spikes represent 95% confidence intervals. Turnover and stayer marginal effects estimated using model 1 with turnover on the right side of the equation. New-to-school marginal effect estimated using model 2 with new to school on the right side of the equation. Tested and untested marginal effects estimated using model 3 with tested on right side of the equation.

schools taking the specified pathway, and spikes represent 95% confidence intervals around those estimates. Panel A draws from the models in Column 1 of Table 2 (value-added scores), while Panel B draws from the models in Column 1 of Table 3 (evaluation score on Standard 1, teacher leadership).

Evidence of strategic staffing with respect to dismissal, retention, and hiring would be evident in the pattern of estimates within locale among the markers in the left panel. For example, the marginal predictions for teacher value-added scores in non-rural schools are lowest among teachers who turned over (approximately -1.2), slightly higher among teachers who stayed (approximately -0.65), and higher still among new-to-school teachers (approximately -0.3 – very close to the state average). While the overlapping confidence intervals show the differences between these teachers in non-rural schools are not statistically significant in the value-added models, Panel B suggests that new-to-school teachers in non-rural schools had significantly higher evaluation scores than those who turned over in the same schools (3.5 relative to 3.2 on the 5-point scale). This pattern suggests that non-rural schools were hiring (and possibly retaining) more effective teachers than they were losing. The pattern is less straightforward in rural schools. Turnovers and stayers were similarly effective on both measures, while new-to-school teachers were descriptively higher on value-added and lower on teacher leadership. The wide confidence interval on the value-added estimate again highlights the noisiness of the value-added measure for new-to-school teachers, who are less likely to have a lagged value-added score because they are more likely to be new to teaching.

Evidence of strategic staffing with regard to teacher assignment would show as a difference between the two sets of markers in the right panel. If schools were assigning their more effective teachers to tested subjects, then the open circles representing value-added for untested teachers would be higher than the triangles representing value-added for untested teachers. Teacher value-added (Panel A) is similar for both sets of teachers in both rural and non-rural schools, though the estimates for untested teachers are particularly noisy since few untested teachers have lagged value-added scores. Panel B provides a more relevant comparison because it is based on evaluation scores that most teachers do have. Specifically, it suggests that the teachers assigned to tested grades and subjects were descriptively more effective than the teachers assigned to untested grades and subjects, though the overlapping confidence intervals show that the estimates are somewhat noisy and the differences are not statistically significant. While this graph focuses on the teacher leadership score, marginal estimates and patterns for the facilitation of learning score and the median evaluation score are very similar. Taken together, the mobility and teacher assignment estimates provide some descriptive evidence that the compositional effect of turnover was positive in non-rural turnaround schools but not in non-rural turnaround schools. By contrast, teacher assignment in rural schools may have nudged up the effectiveness of teachers in tested grades and subjects rather than across the full school.

Table 4 provides the results for teacher experience, with Column 1 operationalizing experience as number of years and Column 2 as a dichotomous indicator for novice teachers. I do not find evidence of differential staffing strategies related to experience for rural and non-rural schools. In both rural and non-rural schools, teachers who turned over (Panel A) and teachers who were new to school (Panel B) had fewer years of experience and were more likely to be novice. Comparing the turnover and new-to-school estimates, it is clear that in the year of CSI identification, these schools lost less experienced teachers than they retained, and then filled those vacancies with even less experienced teachers in the first year of CSI status. Teachers of tested subjects also had fewer years of experience.

Heterogeneity by School Level

In terms of NCEES teacher evaluation scores, I find similar patterns across school levels with respect to teacher turnover. In elementary, middle, and high schools, the teachers who turned over were less effective than the teachers who stayed and there were few detectable differences between rural and non-rural schools. The school-level models also show that the overall finding that new-to-school teachers in rural schools were less effective than new-to-school teachers in non-rural schools appears to be driven by middle and high schools. In particular, new-to-school teachers in rural middle schools were about half a standard deviation less effective on NCEES standards 1 and 4 than new-to-school teachers in non-rural middle schools. The difference was even larger in high schools, with new-to-school teachers in rural schools a full standard deviation less effective on these measures.[5] Together, the school level findings on teacher turnover and new-to-school teachers suggest the compositional effect of turnover was more negative in rural than non-rural middle and high schools in particular.

In the school-level models examining teaching assignment, I find the propensity to assign more effective teachers (as measured by NCEES) to tested subjects is concentrated in middle schools. Specifically, the difference between teachers of tested and untested subjects is about one-fourth to one-third of a standard deviation larger in rural than non-rural middle schools.[6] Put another way, rural middle schools are more likely to assign their more effective teachers to tested subjects than non-rural middle schools. There was no detectable difference between rural and non-rural elementary and high schools.

I provide a snapshot of these differences using NCEES standard 1 in Figure 2. Specifically, Figure 2 displays the coefficient estimates and 95% confidence intervals for Equations 1, 2, and 3, respectively, estimated separately for elementary, middle, and high schools. The coefficient estimates denoted by the circular markers represent the difference in differences between turnovers and stayers in rural compared with non-rural schools. The square markers show the rural-non-rural difference for new-to-school vs. returning teachers, and the triangular markers show the rural-non-rural difference for tested vs. untested teachers. The figure highlights two main takeaways. First, the square markers show that in middle and high schools, the difference in effectiveness between new and returning teachers was more negative in rural than non-rural schools. Second, the triangular markers show that in middle schools in particular, the difference in effectiveness between teachers in tested vs. untested grades and subjects was more positive in rural than non-rural schools. In other words, in certain school levels, non-rural schools were more likely to recruit more effective teachers while rural schools were more likely to assign their more effective teachers to tested subjects.

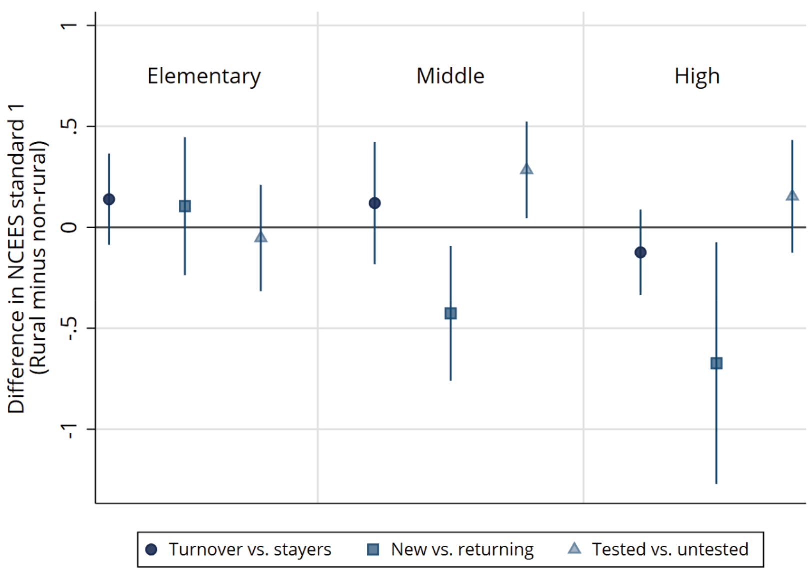

Figure 2. Estimated differences in teacher effectiveness by mobility pathway and school level using NCEES evaluation standard 1 (teacher leadership). Note: Markers represent coefficient estimates for evaluation score on NCEES standard 1 on interaction between rural indicator and turnover (circles), new to school (squares), and tested (triangles) in separate models by school level. Spikes represent 95% confidence intervals. Estimates above the zero line suggest that the difference in effectiveness between groups (i.e., turnovers vs. stayers, new vs. returning, and tested vs. untested teachers) was more positive for rural schools than non-rural schools. Estimates below the zero line suggest that the difference in effectiveness between groups was more negative for rural than nonrural schools.

As with the overall models examining differences in turnover, new-to-school teachers, and course assignment by teaching experience, I do not find meaningful differences in any of the three school levels. These results are provided in appendix Table A1 (NCEES) and Table A2 (experience).

Discussion

This analysis shows that teachers in rural turnaround schools are less effective as measured by value added than teachers in non-rural turnaround schools. It also provides evidence that while non-rural turnaround schools may make personnel decisions to increase the overall effectiveness of the school workforce, rural turnaround schools' personnel decisions do not shift the distribution of teacher effectiveness in their buildings. In fact, these findings point to the possibility of a negative compositional effect of teacher turnover in rural low-performing schools middle and high schools. Any strategic staffing practices in rural schools appear to be targeted at teachers who choose to stay in the building; in particular, rural schools may assign more effective teachers to tested grades and subjects than they assign to untested grades and subjects. These practices may be most prevalent in middle schools. Taken together, these findings suggest that the teacher labor markets for rural turnaround schools may be less robust than for non-rural turnaround schools. These findings are consistent with prior research showing that teacher shortages are more pronounced in rural schools, which tend to have more vacancies and rely more on emergency certification than schools in other locales (Goldhaber & Gratz, 2021; Goldhaber et al., 2020; Nguyen, 2020). This study augments the literature on teacher shortages in rural schools through its explicit focus on low-performing rural schools. The findings are also consistent with the literature on strategic course assignment practices in response to accountability policies (Chingos & West, 2011; Grissom et al., 2017). In particular, the practice of assigning more effective teachers to tested subjects is consistent with documented strategic staffing strategies in accountability contexts.

I highlight three limitations to this analysis. First, these findings are based on purely descriptive analyses from a single year of implementation data (combined with a pre-year of effectiveness data and a post-year of teacher staffing data) and should not be interpreted as causal or as CSI inducing schools to engage in strategic staffing practices. Another limitation to this analysis is the short panel of time observed, as CSI was implemented for just one year before testing were suspended due to COVID-19. Given the short time period, there may be insufficient time to explore how districts are employing strategic staffing over time. Finally, due to the nature of administrative data, I cannot observe whether teachers who leave are dismissed, coached out, or leave voluntarily. In reality, the findings likely reflect a combination of these mechanisms.

The labor market differences for rural and non-rural schools raise two implications for policy and practice related to school turnaround. First, turnaround under ESSA calls for a tailored approach to improvement that accounts for each low-performing school's unique context. To successfully reform low-performing rural schools, the comprehensive planning process will need to consider the local labor market and to apply strategies that improve teaching quality given the labor market constraints. Strategies to support, develop, and retain teachers already on staff may be more efficacious than tactics involving the replacement of less effective teachers because available replacement teachers in rural areas are not more effective than departing teachers. One-size-fits-all approaches to turnaround such as dismissing large swaths of the teacher workforce are unlikely to generalize to rural contexts. Second, a laser focus on improving teacher quality through hiring and assignment in a limited labor market may undermine the stability of turnaround if schools adapt by assigning their least effective teachers to untested early grades.

Low-performing rural and non-rural schools under ESSA have thus far engaged in different strategies with respect to improving teacher effectiveness. Ineffective teachers (i.e., teachers who did not meet expected value-added growth) in rural schools were less likely to turn over than ineffective teachers in non-rural schools in the first year of CSI implementation, and rural turnaround schools in general did not engage in staffing strategies that would meaningfully improve the average effectiveness of their teacher workforce. Instead, rural turnaround schools may have taken steps to maximize the effectiveness of teachers in tested classrooms – though the extent to which this occurred appears to be small and most prevalent in middle schools. It is unclear from the available data whether rural schools may have undertaken other strategies to improve teacher effectiveness, such as professional development or coaching. Also unclear is whether the different strategies employed by rural and non-rural schools were intentional decisions drawing on local context or ad hoc developments driven by the realities of local labor markets. What *is* clear is that the labor markets are different across contexts and school turnaround strategies that do not account for the unique factors associated with rural schools are unlikely to be successful. While low-performing schools in more urban contexts may be able to improve teacher quality by dismissing their least effective teachers, low-performing rural schools are unlikely to find success with such an approach. In fact, to the extent that teacher turnover harms student achievement, dismissing less effective teachers and replacing them with teachers of similar effectiveness levels would reduce student achievement and undermine turnaround. Thus, effectual rural school turnaround may need to improve teacher quality by developing and retaining teachers to maximize the effectiveness of the staff they have. Some research suggests schools can improve recruitment and retention by improving teacher working conditions (Viano et al., 2020), or offering merit or hazard pay (Pham et al., 2021), though there is little evidence on what works specifically in rural low-performing schools and the evidence on merit pay is mixed (Clotfelter et al., 2008a, 2008b).

The findings in this study come from the pre-COVID-19 period. New staffing challenges have emerged nationwide as a result of the COVID-19 pandemic, and given prior evidence on teacher shortages, these challenges are likely exacerbated in low-performing schools and rural low-performing schools in particular. Future research ought to examine differential staffing patterns in rural and non-rural low-performing schools in the years immediately following the pandemic to better understand not just the extent to which the pandemic undermined staffing in low-performing schools but also the ways in which its effects may have varied by school urbanicity. As schools settle into their new normal in coming years, future research can also take a more causal approach to understanding not just whether staffing patterns are different in rural low-performing schools but whether a low-performing designation induces differential responses by school urbanicity. Finally, at a smaller scale, future research could seek out and study interventions aimed at improving teacher quality in low-performing rural schools in particular. Such research will be informative for policy aimed at turning around the lowest performing schools, especially in states such as North Carolina where many of the lowest performing schools are located in rural areas.

Endnotes

1. I exclude 13 charter schools and two lab schools because their autonomy to hire and fire educators is different from in traditional public schools.
2. Thirty-two schools were identified based on graduation rates alone, five were in the bottom 5% in school performance and graduation rates below 66.7%, and the remaining 57 were identified by their school performance score only.
3. The U.S. Census Bureau has four high-level urbanicity codes: urban, suburban, town, and rural. I collapse the first three categories into a single non-rural category because there are very few CSI schools (about 10% total) in suburban and town locales in North Carolina.
4. A similar set of models including school-level covariates find similar patterns and no meaningful gain in precision. I therefore present the models without covariates for simplicity.
5. These estimates were all significant at the $p < 0.05$ level.
6. These estimates were significant at the $p < 0.05$ level.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant Institute of Education Sciences (IES) R305E150017 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education

ORCID

Erica Harbatkin http://orcid.org/0000-0001-8304-2502

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135. https://doi.org/10.1086/508733

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis*, *39*(1), 54–76. https://doi.org/10.3102/0162373716663646

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034–1037. https://doi.org/10.1126/science.1207998

Ayers, J. (2011). Center for American Progress. https://eric.ed.gov/?id=ED535987

Balch, R., & Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review*, *44*, 114–125. https://doi.org/10.1016/j.econedurev.2014.11.002

Barnum, M. (2021, September 23). Teacher vacancies up as year begins. *Chalkbeat*. https://www.chalkbeat.org/2021/9/23/22689774/teacher-vacancies-shortages-covid

Barrow, L. H., & Burchett, B. M. (2001). Needs of Missouri rural secondary science teachers. *The Rural Educator*, *22*(2), 14–19.

Bastian, K. C., & Marks, J. T. (2017). Connecting teacher preparation to teacher induction: Outcomes for beginning teachers in a university-based support program in low-performing schools. *American Educational Research Journal*, *54*(2), 360–394. https://doi.org/10.3102/0002831217690517

Berry, A. B., & Gravelle, M. (2013). The benefits and challenges of special education positions in rural settings: Listening to the teachers. *The Rural Educator*, *34*(2). https://doi.org/10.35608/ruraled.v34i2.400

Bleiberg, J., & Harbatkin, E. (2020). Teacher evaluation reform: A convergence of federal and local forces. *Educational Policy*, *34*(6), 918–952. https://doi.org/10.1177/0895904818802105

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005a). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, *95*(2), 166–171. https://doi.org/10.1257/000282805774669628

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*(1), 113–132. https://doi.org/10.1002/pam.20072

Brownell, M. T., Bishop, A. M., & Sindelar, P. T. (2005). NCLB and the demand for highly qualified teachers: Challenges and solutions for rural schools. *Rural Special Education Quarterly*, *24*(1), 9–15. https://doi.org/10.1177/875687050502400103

Brunner, E., Cowen, J. M., Strunk, K. O., & Drake, S. (2019). Teacher labor market responses to statewide reform: Evidence from Michigan. *Educational Evaluation and Policy Analysis*, *41*(4), 403–425. https://doi.org/10.3102/0162373719858997

Cannata, M. (2010). Understanding the teacher job search process: Espoused preferences and preferences in use. *Teachers College Record*, *112*(12), 2889–2934. https://doi.org/10.1177/016146811011201205 .

Carlson, D., & Lavertu, S. (2018). School improvement grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis*, *40*(3), 0162373718760218. https://doi.org/10.3102/0162373718760218

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

Chingos, M. M., & West, M. R. (2011). Promotion and reassignment in public school districts: How do schools respond to differences in teacher effectiveness? *Economics of Education Review*, *30*(3), 419–433. https://doi.org/10.1016/j.econedurev.2010.12.011

Close, K., Amrein-Beardsley, A., & Collins, C. (2018). National Education Policy Center. https://eric.ed.gov/?id=ED591993

Clotfelter, C., Glennie, E. J., Ladd, H. F., & Vigdor, J. L. (2008a). Teacher bonuses and teacher retention in low-performing schools: Evidence from the North Carolina $1,800 teacher bonus program. *Public Finance Review*, *36*(1), 63–87. https://doi.org/10.1177/1091142106291662

Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008b). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, *92*(5), 1352–1370. https://doi.org/10.1016/j.jpubeco.2007.07.003

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Diaz, R. A. (2004). Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers? *Journal of Policy Analysis and Management*, *23*(2), 251–271. https://doi.org/10.1002/pam.20003

Cohen-Vogel, L. (2011). Staffing to the test: Are today's school personnel practices evidence based? *Educational Evaluation and Policy Analysis*, *33*(4), 483–505. https://doi.org/10.3102/0162373711419845

Collins, T. (1999). *Attracting and retaining teachers in rural areas. ERIC digest.* https://eric.ed.gov/?id=ED438152

Cowen, J. M., Butler, J. S., Fowles, J., Streams, M. E., & Toma, E. F. (2012). Teacher retention in Appalachian schools: Evidence from Kentucky. *Economics of Education Review*, *31*(4), 431–441. https://doi.org/10.1016/j.econedurev.2011.12.005

Cullen, J. B., Koedel, C., & Parsons, E. (2019). The compositional effect of rigorous teacher evaluation on workforce quality. *Education Finance and Policy*, *16*(1), 7–41. https://doi.org/10.1162/edfp_a_00292

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (Working Paper No. 17990). National Bureau of Economic Research. https://doi.org/10.3386/w17990

Dee, T., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267–297. https://doi.org/10.1002/pam.21818

Early, D. M., Maxwell, K. L., Ponder, B. D., & Pan, Y. (2017). Improving teacher-child interactions: A randomized controlled trial of making the most of classroom interactions and my teaching partner professional development models. *Early Childhood Research Quarterly*, *38*, 57–70. https://doi.org/10.1016/j.ecresq.2016.08.005

Fuller, S. C., & Ladd, H. F. (2013). School-based accountability and the distribution of teacher quality across grades in elementary school. *Education Finance and Policy*, *8*(4), 528–559. https://doi.org/10.1162/EDFP_a_00112

Goldhaber, D., & Gratz, T. (2021). *School district staffing challenges in a rapidly recovering economy*. CEDR Flash Brief No. 11082021–1. The Center for Education Data & Research, University of Washington. https://4f730634-e809-4181-92d7-ed1dc705414a.filesusr.com/ugd/1394b9_7709c1ab926247469c2aa9c076b977bc.pdf

Goldhaber, D., & Hansen, M. (2010). Using performance on the job to inform teacher tenure decisions. *American Economic Review*, *100*(2), 250–255. https://doi.org/10.1257/aer.100.2.250

Goldhaber, D., Strunk, K. O., Brown, N., Naito, N., & Wolff, M. (2020). Teacher staffing challenges in California: Examining the uniqueness of rural school districts. *AERA Open*, *6*(3), 2332858420951833. https://doi.org/10.1177/2332858420951833

Goldring, E., Grissom, J. A., Rubin, M., Neumerski, C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher*, *44*(2), 96–104. https://doi.org/10.3102/0013189X15575031

Grissom, J. A., Kalogrides, D., & Loeb, S. (2017). Strategic staffing? How performance pressures affect the distribution of teachers within schools and resulting student achievement. *American Educational Research Journal*, *54*(6), 1079–1116. https://doi.org/10.3102/0002831217716301

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

Hammer, P. C., Hughes, G., McClure, C., Reeves, C., & Salgado, D. (2005). Appalachia Educational Laboratory at Edvantia (NJ1). https://eric.ed.gov/?id=ED489143

Heissel, J. A., & Ladd, H. F. (2018). School turnaround in North Carolina: A regression discontinuity analysis. *Economics of Education Review*, *62*, 302–320. https://doi.org/10.1016/j.econedurev.2017.08.001

Henry, G. T., & Guthrie, J. E. (2019) The Effects of Race to the Top School Turnaround in North Carolina . (EdWorkingPaper: 19-107). Retrieved from Annenberg Institute at Brown University: 10.26300/w488-pf83

Henry, G. T., Guthrie, J. E., & Townsend, L. W. (2015). Outcomes and impacts of North Carolina's initiative to turn around the lowest-achieving schools. http://cerenc.org/wp-content/uploads/2015/09/ES-FINAL-Final-DST-Report-9-3-15.pdf

Henry, G. T., & Harbatkin, E. (2020). The next generation of state reforms to improve their lowest performing schools: An evaluation of North Carolina's school transformation intervention. *Journal of Research on Educational Effectiveness*, *13*(4), 702–730. https://doi.org/10.1080/19345747.2020.1814464

Henry, G. T., McNeill, S. M., & Harbatkin, E. (2021). Accountability-Driven School Reform: Are There Unintended Effects on Younger Children in Untested Grades? (EdWorkingPaper: 19-66). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/4b08-h585

Henry, G. T., Pham, L. D., Kho, A., & Zimmer, R. (2020). Peeking into the black box of school turnaround: A formal test of mediators and suppressors. *Educational Evaluation and Policy Analysis*, *42*(2), 232–256. https://doi.org/10.3102/0162373720908600

Howell, W. G., & Magazinnik, A. (2017). Presidential prescriptions for state policy: Obama's Race to the Top initiative. *Journal of Policy Analysis and Management*, *36*(3), 502–531. https://doi.org/10.1002/pam.21986

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, *38*(3), 499. https://doi.org/10.3102/00028312038003499

Ingersoll, R. M. (2003). *Is there really a teacher shortage? A research report* (September; pp. 1–32). University of Pennsylvania.

James, J., & Wyckoff, J. H. (2020). Teacher evaluation and teacher turnover in equilibrium: Evidence from DC public schools. *AERA Open*, *6*(2), 2332858420932235. https://doi.org/10.1177/2332858420932235

Jimerson, L. (2005). Special challenges of the "No Child Left Behind" Act for rural schools and districts. *Rural Educator*, *26*(3), 1–4. https://eric.ed.gov/?id=EJ783827 .

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment*. MET Project. Bill & Melinda Gates Foundation.

Kaufman, J. H., & Diliberti, M. K. (2021). *Divergent and inequitable teaching and learning pathways during (and perhaps beyond) the pandemic: Key findings from the American educator panels spring 2021 COVID-19 surveys*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA168-6.html

Kraft, M. A., Papay, J. P., & Chi, O. L. (2020). Teacher skill development: Evidence from performance ratings by principals. *Journal of Policy Analysis and Management*, *39*(2), 315–347. https://doi.org/10.1002/pam.22193

Loeb, S., Miller, L. C., & Wyckoff, J. (2015). Performance screens for school improvement: The case of teacher tenure reform in New York City. *Educational Researcher*, *44*(4), 199–212. https://doi.org/10.3102/0013189X15584773

Loeb, S., & Reininger, M. (2004). *Public policy and teacher labor markets. What we know and why it matters*. Education Policy Center.

Malen, B., Croninger, R., Muncey, D., & Redmond-Jones, D. (2002). Reconstituting schools: "Testing" the "theory of action". *Educational Evaluation and Policy Analysis*, *24*(2), 113–132. https://doi.org/10.3102/01623737024002113

Malen, B., & Rice, J. K. (2016). School reconstitution as a turnaround strategy: An analysis of the evidence. In W. Matthis, & T. Trujillo (Eds.), *Learning from the Federal Market-Based Reforms: Lessons for the Every Student Succeeds Act (ESSA)* (pp. 99–125). Charlotte, NC: Information Age Publishing.

Maranto, R. (2013). How do we get them on the farm? Efforts to improve rural teacher recruitment and retention in Arkansas. *Rural Educator*, *34*(1). https://doi.org/10.35608/ruraled.v34i1.406

McClure, C., & Reeves, C. (2004). *Rural teacher recruitment and retention review of the research and practice literature*. Appalachia Educational Laboratory. https://eric.ed.gov/?id=ED484967

Miller, L. C. (2012). Situating the rural teacher labor market in the broader context: A descriptive analysis of the market dynamics in New York State. *Journal of Research in Rural Education (Online)*, *27*(13), 1–31. http://jrre.psu.edu/articles/27-13.pdf

Mobra, T. J., & Hamlin, D. E. (2020). Emergency certified teachers' motivations for entering the teaching profession: Evidence from Oklahoma. *Education Policy Analysis Archives*, *28*, 109. https://doi.org/10.14507/epaa.28.5295

Nguyen, T. D. (2020). Examining the teacher labor market in different rural contexts: Variations by urbanicity and rural states. *AERA Open*, *6*(4), 2332858420966336. https://doi.org/10.1177/2332858420966336

North Carolina State Board of Education, Department of Public Instruction. (2020). *Report to the North Carolina General Assembly: Low-performing districts and schools, improvement planning, and statewide support: 2019-20 school year* (Report to the Joint Legislative Education Oversight Committee on Low Performing Districts and Schools No. 38). Public Schools of North Carolina. https://simbli.eboardsolutions.com/Meetings/Attachment.aspx?S=10399&AID=235414&MID=7758

Papay, J., & Hannon, M. (2018, November 8). *The effects of school turnaround strategies in Massachusetts*. *2018 APPAM Fall Research Conference: Evidence for Action: Encouraging Innovation and Improvement*, Washington, D.C. https://appam.confex.com/appam/2018/webprogram/Paper26237.html

Pham, L. D., Nguyen, T. D., & Springer, M. G. (2021). Teacher merit pay: A meta-analysis. *American Educational Research Journal*, *58*(3), 0002831220905580. https://doi.org/10.3102/0002831220905580

Player, D. (2015). *The supply and demand for rural teachers*. Rural Opportunities Consortium of Idaho.

Proffit, A. C., Sale, R. P., Alexander, A. E., & Andrews, R. S. (2004). The Appalachian model teaching consortium: Preparing teachers for rural Appalachia. *Rural Educator*, *26*(1), 24–29. https://doi.org/10.35608/ruraled.v26i1.518

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. https://doi.org/10.3102/0162373711420864

Rice, J. K., & Malen, B. (2003). The human costs of education reform: The case of school reconstitution. *Educational Administration Quarterly*, *39*(5), 635–666. https://doi.org/10.1177/0013161X03257298

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247–252. https://doi.org/10.1257/0002828041302244

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36. https://doi.org/10.3102/0002831212463813

Rosenberg, L., Christianson, M. D., & Hague Angus, M. (2015). Improvement efforts in rural schools: Experiences of nine schools receiving school improvement grants. *Peabody Journal of Education*, *90*(2), 194–210. https://doi.org/10.1080/0161956X.2015.1022109

Shuls, J. (2018). School Finance in Rural America. In M. Q. McShane, & A. Smarick (Eds.), *No Longer Forgotten: The Triumphs and Struggles of Rural Education in America* (pp. 99–116). Lanham, Maryland: Rowman & Littlefield Publishers.

Stoddard, C., & Toma, E. F. (2021). Introduction to special topic: Rural education finance and policy. *AERA Open*, *7*, 23328584211011610. https://doi.org/10.1177/23328584211011607

Strunk, K. O., Marsh, J. A., Hashim, A. K., Bush-Mecenas, S., & Weinstein, T. (2016). The impact of turnaround reform on student outcomes: Evidence and insights from the Los Angeles unified school district. *Education Finance and Policy*, *11*(3), 251–282. https://doi.org/10.1162/EDFP_a_00188

Sun, M., Penner, E. K., & Loeb, S. (2017). Resource- and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal*, *54*(4), 607–643. https://doi.org/10.3102/0002831217695790

Thompson, C. L., Brown, K. M., Townsend, L. W., Henry, G. T., & Fortner, C. K. (2011). *Turning Around North Carolina's Lowest Achieving Schools (2006-2010)*. Consortium for Educational Research and Evaluation–North Carolina. https://publicpolicy.unc.edu/wp-content/uploads/sites/107/2015/07/Turning-Around-NCs-Lowest-Achieving-Schools-2006-20101.pdf

Viano, S., Pham, L. D., Henry, G. T., Kho, A., & Zimmer, R. (2020). What teachers want: School factors predicting teachers' decisions to work in low-performing schools. *American Educational Research Journal*, *58*(1), 201–233. https://doi.org/10.3102/0002831220930199

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness* (*Second Edition*). New Teacher Project. http://eric.ed.gov/?id=ED515656

Wright, S. P., White, J. T., Sanders, W. L., & Rivers, J. C. (2010). *SAS EVAAS statistical models*. SAS Institute.

Appendices

Table A1. Teacher NCEES.

	ES (1) Standard 1	(2) Standard 4	(3) NCEES median	MS (4) Standard 1	(5) Standard 4	(6) NCEES median	HS (7) Standard 1	(8) Standard 4	(9) NCEES median
Panel A – Teacher turnover (2018-19)									
Rural	0.067	0.053	0.090	0.108	0.214^{+}	0.255^{+}	0.367**	0.302*	0.346*
	(0.083)	(0.088)	(0.082)	(0.108)	(0.121)	(0.130)	(0.126)	(0.139)	(0.157)
Teacher turnover	-0.169^{+}	-0.256**	-0.181**	-0.225*	-0.072	-0.048	-0.130*	-0.134^{+}	-0.154**
	(0.083)	(0.083)	(0.059)	(0.091)	(0.085)	(0.095)	(0.059)	(0.072)	(0.051)
Rural X Teacher turnover	0.139	0.259^{+}	0.097	0.120	0.012	-0.012	-0.124	-0.090	-0.028
	(0.115)	(0.138)	(0.147)	(0.154)	(0.155)	(0.164)	(0.108)	(0.121)	(0.091)
Observations	526	526	386	505	505	367	682	682	423
Panel B – New-to-school teachers (2019-20)									
Rural	-0.080	-0.049	-0.050	0.029	0.070	0.001	0.150	0.233+	0.104
	(0.088)	(0.101)	(0.076)	(0.123)	(0.117)	(0.120)	(0.128)	(0.129)	(0.135)
New to school	-0.081	0.011	-0.071	0.149	0.213+	-0.138	0.064	0.222*	0.173
	(0.124)	(0.090)	(0.122)	(0.129)	(0.106)	(0.103)	(0.124)	(0.103)	(0.114)
Rural X New to school	0.105	-0.019	-0.033	-0.426*	-0.338*	-0.057	-0.673*	-0.779*	-0.461+
	(0.174)	(0.186)	(0.139)	(0.170)	(0.156)	(0.159)	(0.305)	(0.344)	(0.265)
Observations	548	548	390	529	529	363	633	633	394
Panel C – Tested teacher (2018-19)									
Rural	0.129	0.130	0.145	0.026	0.106	0.177	0.301*	0.285+	0.352*
	(0.093)	(0.099)	(0.102)	(0.142)	(0.142)	(0.152)	(0.127)	(0.144)	(0.162)
Teaches tested subject	0.033	-0.038	0.030	-0.025	0.001	0.065	-0.059	-0.004	-0.051
	(0.076)	(0.064)	(0.100)	(0.062)	(0.075)	(0.063)	(0.093)	(0.080)	(0.077)
Rural X Teaches tested subject	-0.053	0.027	-0.057	0.284*	0.256*	0.172	0.153	-0.031	-0.095
	(0.134)	(0.160)	(0.144)	(0.122)	(0.115)	(0.116)	(0.142)	(0.123)	(0.117)
Observations	526	526	386	505	505	367	682	682	423

Standard errors clustered at the school level. All models include school type fixed effects. Most teachers receive scores for Standard 1 and 4 as part of the abbreviated assessment. Teachers with three or fewer years of experience and some other teachers, such as those up for license renewal, receive scores for all five standards as part of the comprehensive assessment. The models predicting median rating restrict the sample to just those teachers with scores on all five standards. NCEES is measured in 2017-18 in Panels A and C and in 2018-19 in Panel B (i.e., the year prior to the pathway). $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A2. Teacher experience.

	ES (1) Teaching experience	(2) Novice teacher	MS (4) Teaching experience	(5) Novice teacher	HS (7) Teaching experience	(8) Novice teacher
Panel A – Teacher turnover						
Rural	1.186	0.010	-1.478	0.098^{+}	2.811*	-0.101^{+}
	(1.064)	(0.049)	(0.897)	(0.055)	(1.154)	(0.058)
Teacher turnover	-0.459	0.100*	-3.034**	0.132^{+}	-0.998	0.092^{+}
	(0.632)	(0.043)	(1.071)	(0.065)	(1.182)	(0.052)
Rural X Teacher turnover	-1.735	0.029	1.808	-0.027	-0.807	0.076
	(1.235)	(0.086)	(1.400)	(0.077)	(1.656)	(0.075)
Observations	756	756	752	752	932	932
Panel B – New-to-school teachers						
Rural	1.368	0.009	-1.074	0.088^{+}	2.851**	-0.100*
	(0.995)	(0.046)	(0.830)	(0.049)	(1.037)	(0.045)
New to school	-3.667**	0.284***	-4.815**	0.319***	-4.115**	0.219**
	(1.128)	(0.065)	(1.351)	(0.074)	(1.273)	(0.063)
Rural X New to school	-1.082	0.007	1.070	-0.124	-4.230*	0.265**
	(1.750)	(0.099)	(1.746)	(0.115)	(1.973)	(0.093)
Observations	782	782	774	774	909	909
Panel C – Tested teachers						
Rural	0.820	-0.002	-0.684	0.091	2.078^{+}	-0.076
	(1.122)	(0.050)	(0.972)	(0.061)	(1.110)	(0.060)
Teaches tested subject	-3.095**	0.095^{+}	-1.004	0.013	-3.760***	0.104^{+}
	(0.846)	(0.052)	(1.011)	(0.061)	(0.839)	(0.053)
Rural X Teaches tested subject	-0.641	0.074	-0.171	-0.015	3.359*	-0.045
	(1.191)	(0.072)	(1.154)	(0.075)	(1.622)	(0.074)
Observations	756	756	752	752	932	932

Standard errors clustered at the school level. All models include school type fixed effects. Novice teachers are those with up to 3 years of teaching experience. Experienced teachers are those with more than 3 years of teaching experience. In Panel A, turnover and experience measured in 2018-19. In Panel B, new-to-school and experience measured in 2019-20. In Panel C, tested classification and experience measured in 2018-19. $^{+}$ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$