Running Head: Federal Policy and the Teacher Labor Market

Federal Policy and the Teacher Labor Market: Exploring the Effects of NCLB on Teacher Turnover

Min Sun

University of Washington

Andrew Saultz

Miami University

Yincheng Ye

Virginia Tech

Suggested Citation: Sun, M., Saultz, A., & Ye, Y. (2014). *Federal Policy and the Teacher Labor Market: Exploring the Effects of NCLB on Teacher Turnover*. Paper presented at the annual meeting of The Association for Education Finance and Policy, San Antonio, TX.

Min Sun (corresponding author) is an Assistant Professor in Quantitative Policy Research in College of Education at the University of Washington. She can be reached via email at: misun@uw.edu or via mail at: 2012 Skagit Lane, M205 Miller Hall (Box 353600), Seattle, WA 98195.

Acknowledgement

Min Sun's contribution to this paper is based upon work supported by the National Science Foundation under Grant No. DRL-1506494. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Federal Policy and the Teacher Labor Market: Exploring the Effects of NCLB School Accountability on Teacher Turnover

Abstract

Anecdotal evidence from the media suggests that accountability pressure increases teacher stress and drives teachers away from teaching, particularly leading to teachers leaving disadvantaged schools that serve a larger proportion of poor and minority students. However, no prior work has systematically examined the changes in the national trends of teacher turnover in response to NCLB school accountability. Drawing on nationally representative samples of schools and teachers from the Schools and Staffing Surveys and Teacher Follow-Up Surveys from 1993-2009, this study applies difference-in-differences approaches to examine the effects of NCLB school accountability on teacher turnover outcomes. We find that NCLB increased the average rate of teachers transferring involuntarily to other schools following school-initiated separations, particularly in the early stage of policy implementation and in disadvantaged schools. More surprisingly, neither voluntarily transferring between schools nor voluntarily leaving the teaching profession has changed in response to NCLB school accountability, even for teachers in disadvantaged schools. In contrast to non-tested counterparts, teachers in tested subjects and grades became less likely to leave the system following school-initiated separations under NCLB school accountability.

Keywords

School Performance Accountability, Federal Policy, Teacher Labor Market

JEL code: J68; I28

Introduction

This study examines the effects of the No Child Left Behind (NCLB) Act of 2001 on teacher turnover. NCLB has dramatically expanded federal influence over U.S. public schools and has been the hallmark of the long-lasting high-stakes accountability reform. Research on the effects of NCLB focuses mostly on its impact on student achievement (e.g., Ballou and Springer 2011; Dee and Jacob 2011; Hemelt 2011; Lee and Reeves 2012; Neal and Schanzenbach 2010; also see a review by Figlio and Loeb 2011). While it is important for policymakers to understand NCLB's impact on this outcome of ultimate interest, it is also important for them to know how NCLB has affected mediating factors in schools, including teachers.

Anecdotal evidence from the media suggests that accountability pressure increases teacher stress (Cavanaugh 2012; Gerson 2007; Toppo 2007). Studies using teacher survey data have shown that when teachers feel their job security is threatened, they experience a reduction in job satisfaction (Finnigan and Gross 2007; Luna and Turner 2001). "Excellent candidates are leaving the career due to [those] political games" (Sunderman, Kim, and Orfield 2005, p. 93). Reback and his colleagues (2014) utilize the differences in accountability pressure across schools and find that teachers in schools on or below the Adequate Yearly Progress (AYP) margin, presumably facing more accountability pressure, are more likely to report concern over their job security related to student test performance in 2003-04, and less likely to plan to teach until retirement than their counterparts teaching in schools with high probabilities of making AYP. In contrast, Grissom et al. (2014) use data from 1994-2008 and find that NCLB accountability has no effect on teachers' general satisfaction with being a teacher at the school, nor on their intent to remain in teaching until retirement.

These mixed findings on the changes in teacher attitudes and career plans provide little guidance on anticipating the impact of NCLB school accountability on teachers' actual employment outcomes. No prior work has applied rigorous methods to nationally representative samples of schools and teachers to examine the changes in national average rates of teachers' turnover after NCLB. It is critical to know the changes in national trends, given that NCLB legislation is a federal policy. This study fills these gaps.

Specifically, this study uses four waves of survey data from the Schools and Staffing Surveys and Teacher Follow-Up Surveys from 1993-95 to 2007-09, and explores three research questions.

1. How did NCLB affect teacher turnover nationwide? We examine the overall impact of NCLB accountability policy on four types of one-year teacher turnover outcomes: transferring to other schools due to teacher-initiated separations (voluntary mobility) or school-initiated separations (involuntary mobility), or leaving the teaching profession altogether due to teacher-initiated separations (voluntary attrition) or school-initiated separations (involuntary attrition).
2. How did the effects of NCLB differ for teachers in tested subjects and grades?
3. How did the effects of NCLB differ for teachers in disadvantaged schools that served a larger proportion of minority or poor students?

This study is important for several reasons. First, high-stakes school accountability continues to be a key component of the U.S. education system. It is useful to understand how this legislation influences the career movements of teachers who have often been regarded as the most important school factor of influencing students' learning (McCaffrey et al. 2003; Rivkin, Hanushek, and Kain 2000; Rowan, Correnti, and Miller 2002; Wright, Horn, and Sanders 1997).

Teacher turnover patterns influence the quality of instruction, sustainability of school improvement, school finance on hiring and firing teachers, and teacher workforce pipeline and distribution across schools. Given the ongoing debates on reauthorizing Elementary and Secondary Education Act (ESEA), both federal and state policymakers are now reflecting on and redesigning accountability policies in public K-12 schools. More rigorous evidence on what works and how it works significantly contributes to current policy debates.

Second, this study distinguishes four types of teacher turnover: voluntary and involuntary mobility and attrition. Often neglected in prior work, this is a very important distinction, because it allows us to test (a) if the school accountability policy drives teachers away because of the job stress and challenging working conditions or (b) if schools become more likely to take staffing actions to separate from low-performing teachers. In addition, the distinction between transfer between schools and exit the teaching profession will reveal if NCLB school accountability influences the redistribution of existing teaching workforce or its overall supply and demand.

Third, we estimate the heterogeneity of the policy impacts for teachers in tested subjects in tested grades and for disadvantaged schools. For example, many claim that NCLB school accountability led to the reallocation of resources to tested subjects and grades (Au 2007; Crocco and Costigan 2007; Dee and Jacob 2011; Zhao 2009). If teachers in tested subjects became more likely to stay in the system in post-NCLB era, this would have supported the hypothesis that schools shift resources to tested subjects to keep those teachers or quickly fill any vacancy. Moreover, NCLB set out to improve student learning in disadvantaged schools. The media has focused heavily on how NCLB has anecdotally changed schools with high poverty and high percentages of racial minorities (Strauss 2015; Viadero 2007). Our study will examine empirically how NCLB influences teacher turnover patterns in these schools.

Lastly, the findings also contribute to the broader literature on how public policies may affect labor market changes (Gibbons and Waldman 1999). The public schooling system, as a large domestic public sector, provides an ideal context to study these labor market issues. NCLB aims to change school practices and teachers' working conditions through performance accountability, which may further alter school administrators' staffing actions and teachers' career decisions. This study documents both direct and indirect evidence on how individuals and organizations engage strategically with policy changes that alter teacher labor market outcomes, and the variation in policy effects for different types of schools and teachers.

To frame our study, we review studies on the impacts of school accountability on teacher turnover and identify gaps in the literature. Next, we describe the data and the sample, and present the analytic strategies for testing our research questions. Lastly, we summarize the main findings and discuss their policy implications on the development of school accountability in the U.S. education system.

School Accountability Impacts on Teacher Turnover

The literature on the impact of accountability policies on teacher turnover contains mixed findings. Loeb and Cuhna (2007) use 1993-94 and 1999-00 waves of the Schools and Staffing Survey (SASS) and the Teacher Follow-up Survey (TFS) and find little impact of states' pre-NCLB accountability efforts on teacher turnover (teachers either moved to other public schools or left the teaching profession), while teachers were more likely to be fired in post-reform years than pre-reform years. Moreover, Shirrell (2013) matches two waves of the Schools and Staffing Survey (1999-01 and 2003-05) with school-level data on the subgroups for which schools nationwide were held accountable during the first two years of NCLB, using the Barnard/Columbia NCLB Database (Reback et al. 2011). He finds that NCLB subgroup-specific

accountability significantly decreased one-year teacher turnover, particularly for Black and Hispanic teachers who taught in a school held accountable for those same racial subgroups. However, this study does not distinguish different types of teacher turnover, such as voluntary or involuntary transferring to another school or leaving the teaching profession altogether.

A number of studies have shown the impact of states' school accountability systems on teacher labor markets. Boyd and his colleagues (2008) use New York State administrative data from 1994-1995 through 2001-2002 to show that the turnover rates of teachers in state-mandated testing grade (i.e., 4th grade) decreased relative to teachers in other elementary grades. Feng, Figlio and Sass (2010)[1] use data collected from Florida schools that experienced accountability "shock" due to changes in state-ordered school grading. Schools that experienced a negative shock (a decreased grade) become less likely to retain their teachers than schools that received no accountability shock. Schools that experienced positive shocks (an increased grade) show a less significant increase in teacher retention than schools experiencing no shock. This accountability effect on teacher attrition is consistent with Clotfelter et al.'s (2004) study in North Carolina and Sims' (2009) study in California. Both studies find that labeling schools as low-performing or failing to meet AYP exacerbates these schools' difficulty in retaining teachers. Hanushek and Rivkin (2010) provide additional evidence through a comparison of teacher transition patterns before and after NCLB in Texas schools. The comparison reveals that the probability of a teacher's transitioning out of a tested grade in a school increased significantly as

[1] Florida had previously graded every school in the state on a scale from "A" to "F" since the summer of 1999, based on proficiency rates in reading, writing and mathematics. In 2002, the state dramatically changed its grading system to both recalibrate the acceptable student proficiency levels for the purposes of school accountability and to introduce student-level changes as an important determinant of school grades. Using student-level micro-data to calculate the school grades that would have occurred absent this change, Feng et al. (2010) demonstrated that over half of all schools in the state experienced an accountability "shock" due to this grading change, with some schools receiving a higher grade than they would otherwise have received (positive shock) and other schools receiving a lower grade than they would have otherwise (negative shock).

the accountability rating[2] declined, even controlling for school demographic characteristics and time trends.

Overall, the existing studies of the impact of school accountability on teacher turnover are limited. For example, few studies that examined school accountability included both pre- and post-NCLB data (e.g., Feng et al. 2010; Shirrell 2013). Instead, most studies collected data either prior to NCLB or on its eve (e.g., Loeb and Cuhna 2007). Further, many of the studies that used both pre- and post-NCLB data were within individual states or districts. Considering the differential implementations of accountability across states (Lee and Reeves 2012) and the variation of local teacher labor markets, the results from a single state can hardly be generalized to the nation (Schneider et al. 2007). Given that NCLB legislation is a federal policy, policymakers would clearly benefit from knowing national trends. More importantly, no study has made the distinction between different types of turnover behaviors under NCLB (e.g., involuntary vs. voluntary transfer between schools or change in profession).

Data and Sample

This study uses nationally representative samples collected by the Schools and Staffing Survey (SASS) and the Teacher Follow-up Survey (TFS), which are sponsored by the National Center for Education Statistics (NCES) and are conducted by the U.S. Census Bureau. This is the most comprehensive data source available on the staffing, occupational, and organizational aspects of nationally representative samples of U.S. elementary and secondary schools and teachers. The SASS uses a stratified probability sample design to select participating schools and teachers in a given year; the following-year TFS sample is a stratified sample from those teachers who answered the SASS teacher survey. To collect the follow-up data, schools are

[2] Texas rated schools on four categories: exemplary, recognized, academically acceptable, and academically unacceptable. The last category is roughly equivalent to not meeting AYP.

mailed a Teacher Status Form at the beginning of the follow-up school year which asks for current information about the previous year's teachers, from which we obtain information on teachers' turnover status. We then obtain information on the reasons for leaving previous schools from the subsequent TFS surveys. The sampling frame allows comparisons of teachers on turnover status by sector (traditional public, public charter, and private), teaching experience, grade, and teacher's race/ethnicity (Tourkin et al. 2010). We merge the SASS and TFS with other federal data sources, including the Common Core of Data (CCD) and the Bureau of Labor Statistics, to obtain more measures.

Four waves of data are used in this study: two waves prior to NCLB (1993–1995, 1999–2001) and two waves after NCLB implementation (2003–2005 and 2007–2009). Many measures collected in the earliest two waves of 1987-88 and 1990-1991 use different metrics from those collected in later years. Therefore, we exclude the first two waves of data. Although the 2012-13 wave is available, using data up to the wave of 2007-09 has the advantage of eliminating alternative explanations on teacher turnover by excluding the influences of the 2008 recession. The Great Recession, which began in 2008, significantly influenced state budgets. The last academic year of data collection for this analysis, the 2008-2009 school year, would not have been effected by recession-related budget cuts, because the school systems would experience such effect mainly starting the 2009-10 academic year (Superfine, Gottlieb, and Smylie 2012), which is beyond the data time frame of this study. Further, the federal program of Race to the Top initiatives began in 2009, which led states to alter teacher policies in a number of ways (Superfine et al. 2012). As a result, using data up to the 2007-09 wave excludes the confounding effects of the two significant nationwide changes in education policy and economic contexts that occurred concurrently with NCLB and could influence teacher labor markets.

Since NCLB provides more specific regulations on grades 3-8 than on other grades, and this legislation offers school-based incentives, our analysis include only public schools that include any of these grades. The analysis excludes charter schools, because the unique mechanisms and structure of charter schools make the modeling of the consequential accountability impact on that teacher labor market complex. We only include full-time classroom teachers.

The analysis is conducted on pooled cross-sectional data at the teacher level and also panel data at the state level. Each wave includes measures on 2,000-3,000 full-time teachers who represent the national population, ranging from about 1.5 million to 2 million teachers. Pooling together these four random samples allows for the development of state-level panel data that contains repetitive observations of the same units of states over time (50 states * 4 waves = 200). Moreover, the SASS sampling frame fortunately allows for generating state-level estimates by grade and school type (Tourkin et al., 2010, p. 58). Table 1 compares the average characteristics of teachers and the schools where they worked prior to and after NCLB implementation.

[TABLE 1 HERE]

Methods

We examine the NCLB impact on four types of one-year teacher turnover outcomes: transferring to other schools in the following year due to teacher-initiated separations (voluntary mobility) or school-initiated separations (involuntary mobility), or leaving the teaching profession altogether due to teacher-initiated separations (voluntary attrition) or school-initiated separations (involuntary attrition). Involuntary mobility or attrition is defined as teachers who reported that the reason for leaving previous year's school was because the teacher's contract

was not renewed at last year's school due to staffing actions (e.g., lay-off, reduction-in-force, reassignment, school closing, or school reorganization).

Our main identification strategy draws on the comparison between a "treatment group" that includes states that did not adopt NCLB-like school accountability policies prior to 2002 and a "comparison group" that adopted such a policy prior to 2002. The premise of this strategy is that NCLB built on the first-generation accountability programs implemented in states such as Texas and North Carolina in the 1990s. The similarities between pre-NCLB state accountability and NCLB suggest that introducing NCLB to states that had already adopted NCLB-like school consequential accountability prior to 2002 would create less of a shock to schools and teachers, and thus be less of a "treatment/intervention" or a close-to-zero treatment. In comparison, treatment effects are anticipated in states that had not previously introduced similar policies.

Although several prior studies analyzing the effects of NCLB have used this strategy (Dee and Jacob 2011; Grissom et al. 2014) and developed the coding for treatment and comparison states, we review all 50 states' pre-NCLB school accountability systems and interrogate prior coding to ensure accuracy. Appendix A details our review in comparison to prior studies, particularly to Hanushek and Raymond (2005) and Dee and Jacob (2011). We agree with prior studies that some states closely resembled NCLB school accountability in that they both publicized school performance and attached the possibility of sanctions to school performance (e.g., takeover, closure, reconstitution, school choice, and replacing school staff) (Dee & Jacob, 2011; Dee, Jacob, and Schwartz, 2013). We code both the existence and timing of states' consequential accountability policies. Table 2 includes our coding of treatment states, which is generally consistent with prior studies, particularly with Dee and Jacob (2011).

[TABLE 2 HERE]

We adopt a difference-in-differences (DID) specification, which essentially estimates the treatment effect by contrasting the difference in the average probability of teacher turnover between pre- and post-NCLB in the treatment group with the difference in the average probability of teacher turnover between pre- and post-NCLB in the comparison group. The function is a multinomial logistic model with a dependent variable in 5 categories (1 = stayer; 2 = voluntary mover; 3 = involuntary mover; 4 = voluntary leaver; and 5 = involuntary leaver). The risk for teacher *i* making a certain type of transition *m* (*m* = 2, 3, 4, 5) relative to staying (*m* = 1, the reference category) in school *j* in state *s* in year *t* is modeled as:

$$\log\left(\frac{\Pr(y_{ijst}=m)}{\Pr(y_{ijst}=1)}\right) = \beta_0 + \beta_1 NCLB_{2004} + \beta_2 NCLB_{2008} + \beta_3(T_s \times NCLB_{2004}) +$$

$$\beta_4(T_s \times NCLB_{2008}) + \beta_5 TEA_{ijst} + \beta_6 SCH_{jst} + \beta_7 STATE_{st} + \mu_s \qquad (1)$$

Where in equation (1), $NCLB_{2004}$ is a dummy indicator for observations in 2004–05 and $NCLB_{2008}$ is a dummy indicator for observations in 2008–09. T_s is a treatment indicator. If state *s* did not have consequential accountability prior to NCLB, it is coded as "1"; it is otherwise coded as "0". β_3 captures the treatment effect in the 2004-05 school year, while β_4 captures the treatment effect in the 2008-09 school year.

We account for several teacher characteristics and school backgrounds that are correlated with teachers' career movements. TEA_{ijst} represents a vector of time-invariant or time-varying teacher covariates, including gender, race/ethnicity, age, being a new teacher who taught three years or less, held a master's degree, state certified, entered the profession through alternative pathways, union status, and taught tested subjects in tested grades. SCH_{jst} includes a vector of school and district covariates, including if the school is a small school (fewer than 200 students) or large school (1,200 students or more), an urban or rural school, or a disadvantaged school if a school satisfies either of two criteria: (1) it served 50 percent or more of students eligible for

free-reduced- price lunch programs (FRPL) (the top 33 percentile points of the distribution of the schools' FRPL rates in 2000); *or* (2) it served 40 percent or more of minority students (the top 33 percentile points of the distribution of the schools' FRPL rates in 2000). The model also includes district instructional expenditure per pupil, indicating the average amount of current expenditures for activities directly associated with the interaction between teachers and students, (e.g., teacher salaries and benefits, supplies, and purchased instructional services). This variable is defined per pupil, in constant 2008 dollars and in natural log.

$STATE_{st}$ includes several time-varying covariates that could influence the year-by-year variability of teacher turnover, including state median household income and state employment population ratio[3]. The variable μ_s represents state fixed effects, which controls for time-invariant variation across states and unobservable state variability. Teacher follow-up final weights are used to make estimates nationally representative and to adjust nonresponse bias (Tourkin et al., 2010). Standard errors are clustered at the state level. Finally, we examine several important threats to causal inference in this study design. The details on robustness/falsification tests are included in the next section.

Results

How Did NCLB School Accountability Affect Teacher Turnover?

The Estimated Effects

Figure 1 compares the means in voluntary and involuntary mobility and attrition rates pre- and post- NCLB, separately for treatment states (the solid line with squares) and comparison

[3] SASS includes multiple measures of working conditions (e.g., teacher classroom autonomy, teachers' earning, professional development, hours per week spent on school-related activities, teacher support from school administration and other colleagues). Although they are identified to be related to teachers' turnover decisions in the literature, we decide not to include them as covariates, because themselves can be influenced by the NCLB policy. Including them as covariates may lead to misidentify the policy treatment effect.

states (the dotted line with dots). The difference in the changes in mean rates of voluntary mobility and attrition from pre- to post-NCLB is small between these two groups, as shown in the upper panel. In contrast, these two groups have more salient differences in the changes of mean rates of involuntary mobility and attrition. As shown in the lower left figure, the involuntary mobility rate increased 1.2% in the treatment group in 2004 (from 1% before NCLB to 2.2% in 2004) and then 0.4% in 2008 (from 1% before NCLB to 1.4% in 2008), while the rate stayed almost unchanged in the comparison group. Moreover, the mean rate of involuntary attrition increased significantly from 0.2% in treatment states prior to NCLB to 1.6% in 2004 and then decreased significantly to roughly 0.1% in 2008. A similar trend occurred in the comparison states as in treatment states: the mean rate increased to 1.1% in 2004 and then decreased to 0.7% in 2008. However, the decrease in involuntary attrition from 2004 to 2008 in treatment states is not as salient as the decrease in control states.

[FIGURE 1 HERE]

Our DID estimation shows similar patterns. As shown in Panel A in Table 3, either the coefficients of voluntary mobility or those of voluntary attrition are close to statistical significance. In contrast, the estimate of the NCLB effect on involuntary mobility in 2004-05, $NCLB_{2004} \times T_s$, is weakly positive (β= 0.832, $p \leq 0.1$). This estimate indicates a larger increase in the likelihood of teachers transferring to another school after school-imitated separations in treatment states than the increase in control states from pre-NCLB to 2004-05. The NCLB effect in 2008-09 on involuntary mobility, as indicated by the estimate of $NCLB_{2008} \times T_s$, is also positive, but not statistically significant. Moreover, the estimate of the involuntary attrition is positive in 2004-05 but statistically insignificant, while the estimate in 2008-09 is significantly negative (β= -2.281, $p \leq 0.01$). This negative estimate indicates that from the pre-NCLB period to 2008, the

increase in the probability of teachers leaving the profession after school-initiated separations in treatment states is less than the increase in comparison states.

In other words, the results of involuntary mobility and attrition suggest that NCLB school accountability policy had a weak effect on increasing teachers' likelihood of transferring to other schools following school-initiated separations in the early year of implementation (i.e., 2004-05), but had a strong effect of retaining teachers in the profession after school-initiated separations in the later year of implementation (i.e., 2008-09). This is a puzzling finding. Is the decrease in the likelihood of leaving the profession in the later year due to fewer teachers leaving or fewer teachers being separated by their previous schools? To assess the changes in the overall school-initiated separations in 2004-05 and 2008-09, we then combine the rates of mobility and attrition and only separate the turnover outcomes by voluntary or involuntary turnover. We apply the equation (1) to the new outcome measure. The estimate of NCLB impacts on involuntary turnover in 2004-05 turns out to be significantly positive ($\beta = 0.751$, $p \leq 0.05$) and the estimate in 2008-09 is not even close to statistically significance ($\beta = -0.066$, $p=0.88$). This evidence suggests that the decrease in school-initiated separations may explain a lot of the decrease in mean rates of involuntary attrition in 2008-09.

Overall, we observe a null effect on the either teachers' voluntarily transferring between schools or voluntarily leaving the teaching profession. The findings contradict the anecdotal claim that NCLB drove teachers away from teaching. Moreover, it seems that NCLB school accountability policy had a stronger effect on school-initiated separations from teachers in the early year of implementation (i.e., 2004-05), particularly involuntarily transferring between schools; but did not increase the likelihood of school separations in later year of implementation (i.e., 2008-09), rather, reduce the chance of separated teachers leaving the profession.

[TABLE 3 HERE]

Robustness and Falsification Analyses

We then conduct multiple robustness or falsification tests of the observed NCLB effects. The key identification assumption of equation (1) is that the trend in teacher turnover from pre- to post-NCLB in the comparison states provides a valid counterfactual for what would have happened in treatment states if NCLB had not been implemented. This assumption would be violated if the changes in teacher turnover prior to NCLB (e.g., 1994 and 2000) were significantly different between treatment and comparison states. To examine this possibility, we test the joint coefficient of ($T_s \times NCLB_{2000}$) and ($T_s \times NCLB_{1994}$). None of the joint coefficients of these four outcome measures shows statistically significant, with p-value between 0.2 and 0.9.

A second threat to the internal validity of the DID model is the plausibility of other concurrent events. That is, the influence on NCLB effects would possibly invalidated if there were unobserved determinants of our outcome measures that varied both contemporaneously with the onset of NCLB school accountability *and* uniquely with respect to treatment status, T_s. As aforementioned, our NCLB effect captures the school accountability effect on teacher labor markets. Other provisions of NCLB, for example, the Highly Qualified Teachers (HQT) provisions, rolled out at the same time as school accountability and could influence schools' actions of laying off teachers who were not highly qualified. And if treatment states implemented the new component of the NCLB—HQT provisions at different levels of rigor from comparison states, the implementation of HQT provisions would bias the estimates. We compare the implementation of HQT using the coding of Loeb and Miller (2006) (see Table A6 in Loeb & Miller, 2006) and do not observe any average differences between these two groups in each of the four HQT requirements: appropriate HQT definitions (chi-square = 0.26, $p<0.878$), public

reporting of HQT data (chi-square = 1.372, $p<0.504$), reporting to the U.S. Secretary of Education on the implementation (chi-square = 0.933, $p<0.624$) and plans to ensure that poor or minority children are not taught by inexperienced, unqualified, or out-of-field teachers at higher rates than are other children (chi-square = 1.531, $p<0.216$). We thus rule out the possibility that the potential differential implementation of HQT would bias the estimates.

There may also be concern about the impact of school closure on the estimated effects on involuntary turnover. Indeed, our data show that school closure occurred both before and after NCLB. The rates varied from 0.39% in 1994-05, 0.43% in 2000-01, 0.63% in 2004-05, and 1.07% in 2008-09. To assess if the increasing trend of school closure is the primary source of the observed effect of NCLB school accountability on involuntary turnover, we exclude teachers whose schools were closed in the years of follow-up surveys. The results, as shown in Panel B in Table 3, are very consistent with those in Panel A.

Additionally, five states adopted a consequential accountability policy in either 2000 or 2001, right before NCLB's passage in 2002 (e.g., Alaska, Delaware, Georgia, Oregon, and Tennessee). They may be weak comparisons because these states' accountability policies might not have been fully implemented before 2002, making them more similar to the treatment states. Moreover, we find it difficult to ascertain, based on available sources, if the state pre-NCLB accountability policy was consequential in Kansas and Virginia. Therefore, we conduct sensitivity analyses by deleting all seven states and re-estimating equation (1). Results are shown in Panel C in Table 3, which are not significantly different from those in Panel A.

We also have concerns about the autocorrelation in the residuals since panel data are employed (Tyler et al. 2000). We use the Durbin-Watson test (D-W) to detect the presence of observable autocorrelation. The D-W statistic of 2 indicates no autocorrelation, while a result

substantially different than 2 indicates either positive or negative serial correlation. In our analysis, the largest value of the D-W statistics is 2.66 (or autocorrelation coefficient t = -0.33). It appears that autocorrelation is not a major concern.

Lastly, we conduct a falsification test using a fake year for NCLB onset and a fake treatment assignment. We use the 1999-00 school year as the onset of NCLB rather than 2002, and randomly assign states into treatment and control categories using random generators to create a fake treatment assignment. Table 4 includes the results that none of the coefficients is statistically significant, which supports the claim that the policy effects identified in Panel A Table 3 are systematic, rather than purely by chance.

[TABLE 4 HERE]

Heterogeneity of NCLB Effects on Subgroups of Teachers and Schools

Teachers who Taught Tested Subjects in Tested Grades

Because states are mandated to test grades 3-8 in math and reading and categorize schools based on the results, it is anticipated that teachers who taught in this grade span and subject areas would feel more accountability pressure than other non-tested teachers. We define a tested group as general teachers in elementary schools who taught grades 3-5, or subject-specific teachers who taught math or reading in grades 3-8. This tested group includes about 36% to 42% of total teachers in the samples across waves. The rest of teachers are defined as in the non-tested group.

Table 5 includes the estimates. The first two panels include the results of applying equation (1) to only tested or non-tested teachers separately, while the last panel includes the results of applying a three-way interaction term of $NCLB \times T_s \times Tested$ to the whole sample to assess if the differences in the coefficients are statistically significant. Overall, we do not observe

any statistically significant differences between these two groups in NCLB school accountability effects on voluntary mobility, involuntary mobility or voluntary attrition. However, we observed differential policy effects on involuntary attrition between these two types of teachers. After school-initiated separations, tested teachers did not increase their risk of leaving the profession at the early stage of NCLB implementation in 2004-05 (β = - 1.47, p >0.1), while non-tested teachers increased their likelihood of leaving the profession (β =1.413, $p\leq$ 0.05). The difference in coefficients is statistically significant, as captured by the interaction effect of $NCLB_{2004}\times T_s\times Tested$ (β =-2.786, $p\leq$ 0.05). Tested teachers continued to be more likely to be retained in 2008 than non-tested counterparts, although the difference is statistically insignificant.

[TABLE 5 HERE]

Disadvantaged Schools

NCLB set out to improve student learning in disadvantaged schools. The media has focused heavily on how NCLB has anecdotally changed schools with high poverty and high percentages of racial minorities (Viadero 2007; Strauss 2015). The wide held belief is that teachers in these schools are under more pressure to meet AYP targets. To test this, we conduct separate analyses for disadvantaged schools and advantaged counterparts using similar methods as those for tested teachers. We define *disadvantaged schools* if a school satisfies either of two criteria: (1) it served 50 percent or more of students eligible for free-reduced- price lunch programs (FRPL) (the top 33 percentile points of the distribution of the schools' FRPL rates in 2000); *or* (2) it served 40 percent or more of minority students (the top 33 percentile points of the distribution of the schools' FRPL rates in 2000)[4]. As contingent upon the change in the

[4] To further confirm the above findings, we separated schools into three groups using flexible thresholds to reflect the increase in the average percentages of FRPL and minority students over years: the advantaged schools (top 33 percentile points of the distributions in a year), the most disadvantaged schools (the bottom 33 percentile points),

nationwide poverty rates and the racial composition of the national population over years, there were about 41% of schools in our sample were categorized as *disadvantaged schools* in 1993-94. This percentage gradually increased over time, to about 57% in 2008-09. Any school that did not meet either criterion for a disadvantaged school is coded as an advantaged school.

As shown in the second panel of Table 6, NCLB school accountability shows no significant effects in advantaged schools in any of the turnover categories. However, it shows a significant effect on the increase of involuntary mobility rates in disadvantaged schools in 2004-05 (β= 1.548, $p\leq0.05$) and a weaker increase in 2008-09 (β= 1.306, $p\leq0.1$) (see the first panel of Table 6). This federal policy also significantly reduced the likelihood of a teacher who was separated by a disadvantaged school to leave the teaching profession (β= -2.732, $p\leq0.01$) in 2008-09. The differences in NCLB effects between disadvantaged schools and advantaged schools are more salient in 2008-09, as indicated by the $NCLB_{2008}\times T_s\times$*Disadvantaged Schools* estimates ($\beta_{involuntary\ mobility}$ = 1.763, $p\leq0.1$; $\beta_{\text{-}involuntary\ attrition}$= -2.59, $p\leq0.1$). Similar inferences of NCLB effects are identified after excluding closure schools[5].

[TABLE 6 HERE]

Discussion

How did NCLB school accountability influence school staffing and teacher turnover outcomes? Our investigation suggests that NCLB school accountability did not change the average national rate of teachers either voluntarily transferring between schools or voluntarily leaving the teaching profession in the following year. This null effect accords with findings about NCLB impacts on other outcome measures about teachers. For example, Grissom et al. (2014) note that NCLB had no measurable effects on teachers' general satisfaction with being a

and the middle. The inferences are generally consistent with those in Table 6 (results are available upon request from the authors).

[5] Results are available upon request from the authors.

teacher at their school, nor on their intent to remain in teaching until retirement. Our results also accord with the Loeb and Cuhna's (2007) findings that pre-NCLB state school accountability had no effects on voluntary teacher turnover, but positive effects on the probability of teachers being separated by their previous schools.

This study extends Loeb and Cuhna's work by further separating school-initiated separations into transferring between schools or leaving teaching profession. We find that NCLB school accountability had a stronger effect on school-initiated separations from teachers in the early stage of implementation (i.e., 2004-05), particularly involuntarily transferring between schools. However, NCLB did not increase the likelihood of school-initiated separations in the later stage of implementation (i.e., 2008-09); particularly, it reduced the chance of separated teachers leaving the profession. Since we control for state economic condition and district instructional expenditure, it is less likely that the differences between 2004-05 and 2008-09 are driven by the national or local economy. We also control for school size to account for the fluctuation of student enrollment over time. Moreover, our robustness check rules out the possibility that the increase in involuntary mobility in 2004-05 was due to school closure. Therefore, one plausible explanation could be that schools experienced a greater shock in the early stage of implementation thus significantly changed their staffing strategy. Such policy shock effect died down in the later stage of implementation.

When we analyze the policy effects separately for subgroups of teachers, we find that under NCLB school accountability, tested teachers were less likely to leave the system following school-initiated separations. This finding may be a result of the tested group gaining demand in the labor market during the NCLB era, while the non-tested teachers might experience the decrease in demand. In this regard, schools were more motivated to fill a vacancy due to school-

initiated separation in a tested area than a non-tested area. This would fit into the argument that schools shift resources from non-tested areas to tested areas, as illustrated in prior studies of narrowing the curriculum (Au 2007; Crocco and Costigan 2007; Zhao 2009) and reallocating the instructional time to tested areas and away from non-test areas (Dee et al. 2013; Reback, Rockoff, and Schwartz 2014).

Additionally, this study analyzes differential policy impacts for disadvantaged schools and advantaged counterparts. While NCLB has little impact in advantaged schools, this policy is estimated to increase teacher involuntary mobility in disadvantaged schools. These results fit into the popular speculation that disadvantaged schools were affected by NCLB to a larger degree than advantaged schools. Disadvantaged schools were more likely to separate from their teachers because (a) they hired a higher percentage of new, poorly credentialed or less effective teachers to begin with, and (b) these types of schools were more pressured to change existing workforce and school practices. In addition, NCLB school accountability has no influences on the changes in teachers' voluntary turnover, even in disadvantage schools. This counters the critiques from many educational voices that charge NCLB with driving teachers away from schools serving the neediest students.

The analyses here have some limitations. Due to the lack of detailed data on the teacher labor market, we could not identify if involuntary mobility resulted from teachers' own efforts to secure new jobs (the supply side) or from school-arranged transfers (the demand side). It is also highly possible that although teachers reported that they had left the previous year's schools voluntarily, school administrators actually encouraged them to leave. If this was common in schools that faced school accountability pressure, this analysis may have underestimated the effects of NCLB school accountability on teachers' involuntary turnover. Because this study

cannot tell if the school-separated teachers were less effective in terms of teachers' ability to improve students' learning gains, whether NCLB achieves the goal of improving the quality of instruction is also unclear. Lastly, the time series has limited observations. The long-term impact of NCLB still remains to be seen.

Despite these limitations, this study's findings can help to anticipate the impacts of NCLB waivers. Across the board, the NCLB waiver program reduces the number of schools identified for interventions, often substantially (Polikoff et al. 2014). The study may also help to project future consequences of the Alexander-Murray bill of Elementary and Secondary Education Act (ESEA) reauthorization released on April 7, 2015. The Alexander-Murray bill also gives more control to states over what their school accountability systems look like, although states systems would still have to consider student achievement and the performance of student subgroups. The decrease in accountability pressure from the federal government may reduce school administrators' motivation to separate from teachers, particularly in underperforming schools. Moreover, because NCLB focused on schools as the unit of analysis, rather than individual teachers, we observed NCLB's impact only on school-initiated separations but a null effect on teacher-initiated career movements. In contrast, the teacher quality initiatives under waivers and reauthorization highlight individual teachers—for example, teacher evaluation under waivers and teacher incentives enshrined in the Alexander-Murray bill. These may possibly generate more significant effects on teacher-initiated turnover than what we observed under NCLB.

These findings add empirical evidence to the idea that employment outcomes are affected deeply by public policies, not determined in an institution-free labor market (Gibbons and Waldman 1999). Federal policy, with NCLB accountability as an example, clarifies the

performance assessment standards for public employers and employees, and uses policy levers to align their values and behaviors. It is important for researchers to continue to explore how organizations and individuals engage strategically with policy and management change, as the national dialogue regarding the future of accountability in public education continues to focus on school accountability and teacher quality.

References

Amrein, Audrey L., and David C. Berliner. 2002. High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10*(18): 1-74.

Au, Wayne. 2007. High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5): 258-267.

Ballou, Dale, and Matthew G. Springer. 2008. *Achievement trade-offs and No Child Left Behind.* Washington D.C.: Urban Institute. Retrieved May 20, 2012, from http://www.urban.org/UploadedPDF/1001309_achievement_tradeoffs.pdf

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2008. The impact of assessment and accountability on teacher recruitment and retention: Are there unintended consequences? *Public Finance Review, 26*(1): 88-111.

Carnoy, Martin, and Susanna Loeb. 2002. Does external accountability affect student outcomes? A cross state analysis. *Educational Evaluation and Policy Analysis, 24:* 305-331.

Cavanaugh, Sean. 2012. Survey: N.C. teachers say high-stakes tests dominate classes. *Education Week,* March 8. Retrieved from http://blogs.edweek.org/

Clotfelter, Charles T., Helen F. Ladd, Jacob L. Vigdor, and Roger Aliaga Diaz. 2004. Do school accountability systems make it more difficult for low performing schools to attract and retain high quality teachers? *Journal of Public Policy Analysis and Management, 23*(2): 251–271.

Crocco, Margaret S., and Arthur T. Costigan. 2007. The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education, 42*(6): 512-535.

Dee, Thomas S., and Brian Jacob. 2011. The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*(3): 418-446.

Dee, Thomas S., Brian Jacob, and Nathaniel Schwartz. 2013. The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis, 35*(2): 252-279.

Feng, Li, David N. Figlio, and Tim Sass. 2010. *School accountability and teacher mobility (working paper 16070).* Cambridge, MA: National Bureau of Economic Research.

Figlio, David N., and Susanna Loeb. 2011. School accountability. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.), *Handbooks in Economics: Economics of Education* (vol. 3, pp. 383-421). The Netherlands: North-Holland.

Finnigan, Kara S., and Betheny Gross. 2007. Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal, 44*(3): 594-629.

Gerson, Jack. 2007. NCLB: Bad for teachers, bad for kids. *United Teacher, 37*(2): 14-32.

Gibbons, Robert, and Michael Waldman. 1999. Careers in organizations: Theory and evidence. In in *Handbook of Labor Economics* (eds. Orley C. Ashenfelter and David Card. Vol. 3) Amsterdam: North-Holland.

Grissom, Jason A., Sean Nicholson-Crotty, and James R. Harrington. 2014. Estimating the effects of No Child Left Behind on teachers' work environments and job attitudes. *Educational Evaluation and Policy Analysis, 36*(4): 417-436.

Hanushek, Eric A., and Margaret E. Raymond. 2005. Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management, 24* (2): 297-327.

Hanushek, Eric A., and Steven G. Rivkin. 2010. The quality and distribution of teachers under the No Child Left Behind Act. *Journal of Economic Perspectives, 24*(3): 133-150

Hemelt, Steven W. 2011. Performance effects of failure to make Adequate Yearly Progress (AYP): Evidence from a regression discontinuity framework. *Economics of Education Review, 30:* 702-723

Lee, Jaekyung, & Todd Reeves. 2012. Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990-2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis, 34*(2): 209-231.

Loeb, Susanna, & Luke Miller. 2006. *A federal foray into teacher certification: Assessing the Highly qualified teacher' provision of NCLB*. Report prepared for The Campaign for Educational Equity at Teachers College, Columbia University Symposium, Examining America's Commitment to Closing Achievement Gaps NCLB and Its Alternatives. Retrieved at: http://cepa.stanford.edu/content/federal-foray-teacher-certification-assessing-highly-qualified-teacher-provision-nclb#sthash.dVOZ0v8f.dpuf

Loeb, Susanna, and Jesse Cuhna. 2007. *Have assessment-based accountability reforms influenced the career decisions of teachers and principals*? A report commissioned by the U.S. Congress as part of Title I, Part E, Section 1503 of the No Child Left Behind Act of 2001. Retrieved from https://cepa.stanford.edu/sites/default/files/Cunha_Accountability_Labor_Decisions.pdf.

Luna, Catherine, and Cara livingstone Turner. 2001. The impact of the MCAS: Teachers talk about high-stakes testing. *The English Journal, 91*(1): 79-87.

McCaffrey, Daniel F., J.R. Lockwood, Daniel M. Koretz, and Laura S. Hamilton. 2003. *Evaluating value-added models for teacher accountability*. Arlington VA: RAND Corporation.

Neal, Derek, and Diane Whitemore Schanzenbach. 2010. Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics, 92*(2): 262-283.

Polikoff, Morgan S., Andrew J. McEachin, Stephani L. Wrabel, and Matthew Duque. 2013. *Educational Researcher, 43*(1): 45-54

Reback, Randall, Jonah Rockoff, Heather Schwartz, and Elizabeth Davidson. 2011. *Barnard/Columbia No Child Left Behind Database, 2002-2003 and 2003-2004*, retrieved from http://www.gsb.columbia.edu/nclb.

Reback, Randall, Jonah Rockoff, Heather Schwartz. 2014. Under pressure: Job security, resource allocation, and productivity in schools under No Child Left Behind. *American Economic Journal: Economic Policy, 6*(3): 207-41.

Rivkin, Steven G., Eric A. Hanushek, & John F. Kain. 2005. Teachers, schools, and academic achievement. *Econometrica, 73*(2): 417-458.

Rowan, Brian, Richard Correnti, and Robert J. Miller. 2002. *What large-scale, survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools*. Philadelphia, PA: Consortium for Policy Research in Education.

Schneider, Barbara, Martin Carnoy, and Jeremy Kilpatrick et al. 2007. *Estimating causal effects: Using experimental and observational designs.* Washington, D.C.: American Educational Research Association.

Shirrell, Matthew A. 2013. *The effects of subgroup-specific accountability on teacher attrition*. Washington, D.C.: Paper presented at the Fall Research Conference of the Association for Public Policy Analysis and Management.

Sims, David P. 2009. *"Going down with the ship?" The effect of school accountability on the distribution of teacher experience in California*. Retrieved from http://www.caldercenter.org/upload/NCLB-paper-Sims-4.pdf

Strauss, Valerie. 2015. No child left behind's test-based policies failed. Will Congress keep them anyway? *The Washington Post*, February 13. Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/wp/2015/02/13/no-child-left-behinds-test-based-policies-failed-will-congress-keep-them-anyway/

Sunderman, Gail L., James S. Kim, and Gary G. Orfield. 2005. *NCLB meets school realities: Lessons from the field*. Thousand Oaks, CA: Corwin.

Superfine, Benjamin M., Jessica J. Gottlieb, and Mark A. Smylie. 2012. The expanding federal role in teacher workforce policy. *Educational Policy, 26* (1): 58-78.

Toppo, Greg. 2007. How Bush education law has changed our schools. *USA Today*, August 1. Retrieved from http://usatoday30.usatoday.com/news/education/2007-01-07-no-child_x.htm.

Tourkin, Steven, Teresa Thomas, Nancy Swaim, and Shawna Cox et al. 2010. *Documentation for the 2007–08 Schools and Staffing Survey* (NCES 2010-332). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

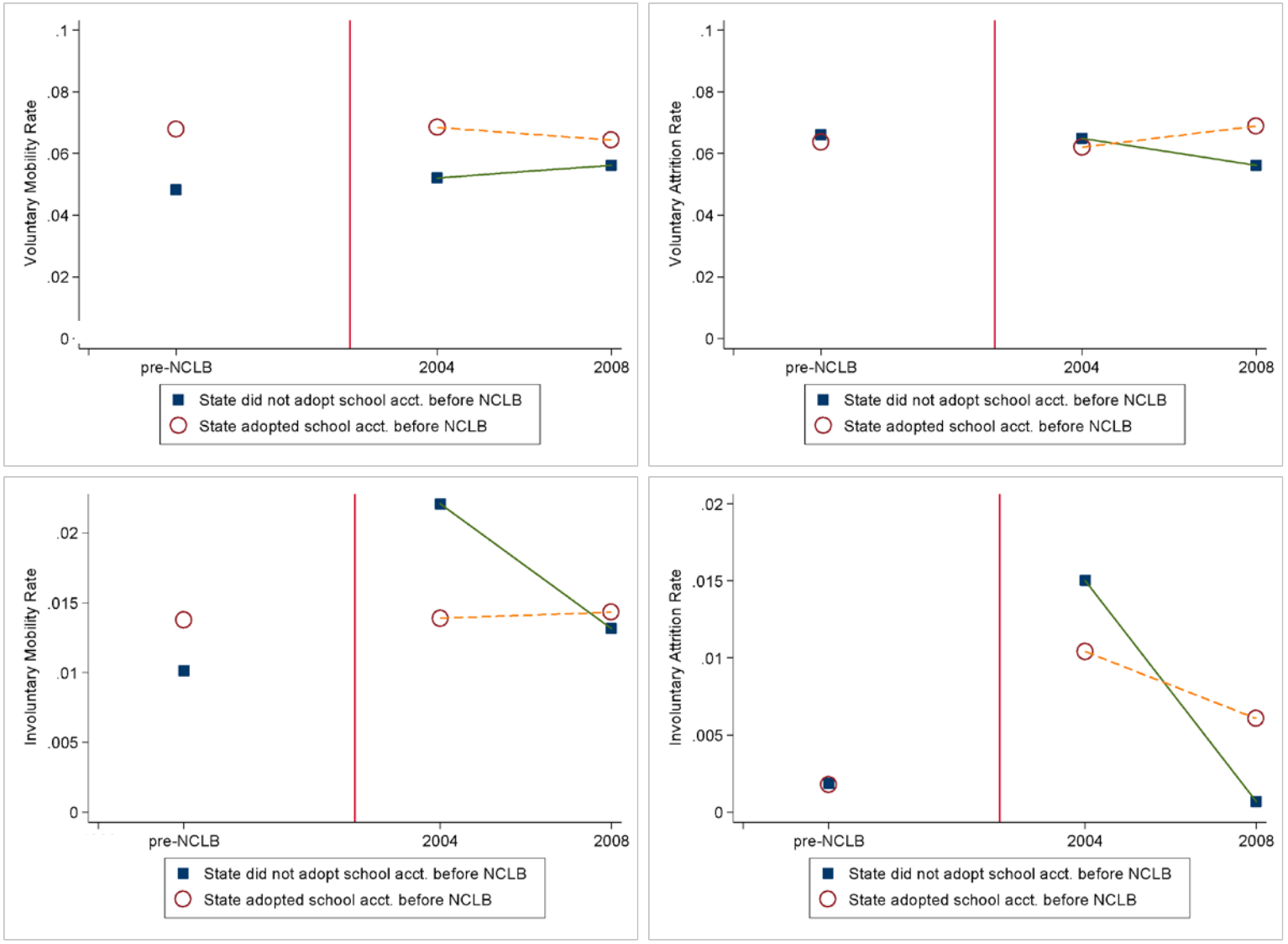
Tyler, John H., Richard J. Murnane, and John B. Willett. 2000. Estimating the labor market signaling value of the GED. The Quarterly *Journal of Economics, 115*(2):431-468.

Viadero, Debra. 2007. No easy answers about NCLB's effect on "poverty gap". *Education Week.* November 14. Retrieved from http://www.edweek.org/ew/articles/2007/11/14/12brookings.h27.html

Wright, S. Paul, Horn, Sandra P., and Willaim L. Sanders. 1997. Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*: 57-67.

Zhao, Yong. 2009. *Catching up or leading the way*. Alexandria, VA: ASCD.

Figure 1. Compare teacher average turnover rates by treatment groups between pre-and post-NCLB years



Note: Two figures on the upper panel are on the same scale (0-0.1) and two figures on the lower panel on the same scale (0-0.02). The squares and dots indicate the means of each treatment group-by-NCLB status cell. Final eacher follow-up weights were used in the calculation of the national average turnover rates.

Table 1. Descriptive statistics

	Pre-NCLB	Post-NCLB	Wald Test
Teacher characteristics			
Pct. Female	0.81	0.82	0.5
	(0.393)	(0.384)	
Ave. age	42.445	42.22	0.35
	(10.308)	(11.46)	
Pct. teachers held a Master's and higher degree	0.455	0.483	2.51
	(0.498)	(0.5)	
Pct. certified teachers	0.963	0.992	39.18***
	(0.19)	(0.088)	
Pct. entered teaching via alternative pathways	0.029	0.086	44.96***
	(0.169)	(0.28)	
Ave. teaching experience	14.528	13.82	4.12*
	(9.753)	(10.204)	
Pct. union membership	0.807	0.803	0.07
	(0.395)	(0.398)	
School contexts			
Pct. small schools	0.096	0.076	5.32*
	(0.294)	(0.265)	
Pct. large schools	0.034	0.044	6.99***
	(0.181)	(0.21)	
Ave. percentage of White students	0.637	0.56	61.97***
	(0.335)	(0.347)	
Ave. percentage of Black students	0.159	0.185	14.51***
	(0.246)	(0.26)	
Ave. percentage of Hispanic students	0.157	0.207	37.22***
	(0.251)	(0.279)	
Pct. rural schools	0.333	0.195	141.35***
	(0.471)	(0.396)	
Pct. urban schools	0.27	0.288	2.15
	(0.448)	(0.453)	
Ave. percentage of students eligible for free- or reduced-price lunch (FRPL)	0.409	0.491	87.63***
	(0.29)	(0.304)	
State			
Ave. instructional expenditure per pupil (in constant 2008 dollars and in natural log)	8.484	8.678	209.45***
	(0.222)	(0.235)	
Ave. state median household income(in constant 2008 dollars and in natural log)	4.705	4.691	10.28**
	(3.882)	(3.898)	
Ave. state employment population ratio	0.644	0.634	5.37*
	(0.042)	(0.041)	
Sample teacher N	4,570	4,780	
Population teacher N	3,272,340	3,958,220	

Notes: Means and its standard deviations are calculated by applying appropriate sampling weights to make these descriptive statistics nationally representative. Standard deviations are included in parentheses.
All sample sizes are rounded to the nearest ten to comply with the restricted-use data agreement with NCES.
† $p \leq 0.1$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

Table 2. Coding of treatment states that had school accountability similar to NCLB prior to 2002

	Our coding		**Dee/Jacob (2011)**		**Hanushek /Raymond (2005)**	**Carnoy /Loeb (2002)**	**Amrien /Berliner (2002)**
State	Year	Treatment	Year	CL acct.			
AL	1997	0	1997	Yes	CL (97)	Strong	4
AK	2001	0	2001	Yes	n/a	None	0
AZ	2002	1	2002	No	n/a	None	0
AR	1999	0	1999	Yes	CL (99)	None	0
CA	1999	0	1999	Yes	CL (99)	Strong	5
CO	2002	1	2002	No	n/a	None	5
CT	1999	0	1999	Yes	CL (93)	Weak	0
DE	2001	0	1998	Yes	CL (98)	None	6
FL	1999	0	1999	Yes	CL (99)	Strong	5
GA	2001	0	2000	Yes	CL (00)	None	1
HI	2002	1	2002	No	n/a	None	0
ID	2002	1	2002	No	n/a	None	0
IL	1992	0	1992	Yes	n/a	Moderate	0
IN	1994	0	1994	Marginal	Report Card (93)	Moderate	4
IA	2002	1	2002	No	n/a	None	0
KS	1995	0	1995	Marginal	Report Card (93)	Weak	0
KY	1995	0	1995	Yes	CL (95)	Strong	4
LA	1999	0	1999	Yes	CL (99)	Moderate	5
ME	2002	1	2002	No	n/a	None	0
MD	1999	0	1999	Yes	CL (99)	Strong	5
MA	1998	0	1998	Yes	CL (98)	Implicit Only	3
MI	1998	0	1998	Yes	CL (98)	Weak	5
MN	2002	1	2002	No	n/a	None	1
MS	2002	1	2002	No	n/a	None	2
MO	2002	1	2002	No	n/a	None	1
MT	2002	1	2002	No	n/a	None	1
NE	2002	1	2002	No	n/a	None	0
NV	1996	0	1996	Yes	CL (96)	Weak	4
NH	2002	1	2002	No	n/a	None	0
NJ	2002	1	2002	No	n/a	None	3
NM	1998	0	1998	Yes	CL (03)	Moderate to strong	5
NY	1998	0	1998	Yes	CL (98)	Strong	4
NC	1996	0	1996	Yes	CL (96)	Strong	6
ND	2002	1	2002	No	n/a	None	0
OH	2002	1	2002	No	n/a	None	5
OK	1996	0	1996	Yes	CL (97)	Weak	2
OR	2000	0	2000	Yes	CL (00)	Weak to	0

						moderate	
PA	2002	1	2002	No	n/a	None	3
RI	1997	0	1997	Yes	CL (97)	Weak	0
SC	1999	0	1999	Yes	CL (99)	Moderate	5
SD	2002	1	2002	No	n/a	None	0
TN	2000	0	2000	Yes	CL (96)	Weak	4
TX	1994	0	1994	Yes	CL (94)	Strong	6
UT	2002	1	2002	No	n/a	None	0
VT	1999	0	1999	Yes	CL (99)	Weak	0
VA	1998	0	1998	Marginal	CL (98)	Weak to moderate	2
WA	2002	1	2002	No	n/a	None	0
WV	1997	0	1997	Yes	CL (98)	Strong	3
WI	1993	0	1993	Marginal	CL (93)	Weak to moderate	0
WY	2002	1	2002	No	n/a	None	0

Notes: “CL” stands for consequential, and “acct.” stands for accountability.

Table 3. The estimated effects of NCLB school accountability on teacher turnover

	Voluntary Mobility	Involuntary Mobility	Voluntary Attrition	Involuntary Attrition
Panel A. Main model				
$NCLB_{2004} \times T_s$	0.012	0.832†	0.046	0.409
	(0.223)	(0.443)	(0.234)	(0.834)
$NCLB_{2008} \times T_s$	0.254	0.301	-0.181	-2.281**
	(0.250)	(0.492)	(0.314)	(0.861)
Sample N		9,270		
Panel B. Excluded schools that were closed next year				
$NCLB_{2004} \times T_s$	0.074	0.793†	-0.206	0.295
	(0.231)	(0.469)	(0.333)	(0.823)
$NCLB_{2008} \times T_s$	0.279	0.124	-0.478	-2.372**
	(0.257)	(0.457)	(0.37)	(0.866)
Sample N		9,010		
Panel C. Excluded marginal states				
$NCLB_{2004} \times T_s$	0.016	0.772†	0.036	0.284
	(0.229)	(0.469)	(0.250)	(0.847)
$NCLB_{2008} \times T_s$	0.22	0.400	-0.048	-2.483**
	(0.252)	(0.495)	(0.324)	(0.878)
Sample N		8,120		

Notes. † $p \leq 0.1$; * $p \leq 0.05$; ** $p \leq 0.01$

(a) This model includes main effects only. Controls include teacher's gender, race/ethnicity, age, being a new teacher who taught three years or less, held a master's degree, state certified, entered the profession through alternative pathways, union status, and taught tested subjects in tested grades. It also includes school-level covariates (e.g., small or large small, disadvantaged schools, urban or rural school, and district instructional expenditure per pupil) and state-level covariates (e.g., state median household income and state employment population ratio).
(b) State fixed effects are included.
(c) Standard errors are clustered at state level and estimates are included in parentheses.
(d) Each model was estimated applying teacher follow-up final weights to make the results nationally representative and adjust for nonresponse bias.
(e) All sample sizes are rounded to the nearest ten due to the restricted-use data agreement with NCES.

Table 4. Falsification tests

	Voluntary Mobility	Involuntary Mobility	Voluntary Attrition	Involuntary Attrition
Fake year of on-site NCLB [(1)]	0.007	0.075	0.111	-0.026
	(0.187)	(0.279)	(0.151)	(0.45)
Sample N		9,270		
Fake $T_s \times NCLB_{2004}$	-0.106	0.189	-0.214	0.495
	(0.147)	(0.289)	(0.171)	(0.666)
Fake $T_s \times NCLB_{2008}$	-0.274	0.328	-0.078	0.451
	(0.205)	(0.385)	(0.310)	(0.388)
Sample N		9,270		

Note: [(1)]the fake year is 1999-2000.

(a) This model includes main effects only. Controls include teacher's gender, race/ethnicity, age, being a new teacher who taught three years or less, held a master's degree, state certified, entered the profession through alternative pathways, union status, and taught tested subjects in tested grades. It also includes school-level covariates (e.g., small or large small, disadvantaged schools, urban or rural school, and district instructional expenditure per pupil) and state-level covariates (e.g., state median household income and state employment population ratio).

(b) State fixed effects are included.

(c) Standard errors are clustered at state level and estimates are included in parentheses.

(d) Each model was estimated applying teacher follow-up final weights to make the results nationally representative and adjust for nonresponse bias.

(e) All sample sizes are rounded to the nearest ten due to the restricted-use data agreement with NCES.

Table 5. NCLB effects on teachers who taught tested subjects and grades

	Voluntary Mobility	Involuntary Mobility	Voluntary Attrition	Involuntary Attrition
Teachers taught tested subjects/grades				
$NCLB_{2004} \times T_s$	-0.207	1.046	-0.226	-1.47
	(0.341)	(0.697)	(0.393)	(1.105)
$NCLB_{2008} \times T_s$	0.523	0.423	-0.459	-3.386*
	(0.361)	(0.638)	(0.29)	(1.534)
Sample N		3,560		
Teachers taught low-stakes subjects/grades				
$NCLB_{2004} \times T_s$	0.15	0.612	0.259	1.413*
	(0.344)	(0.427)	(0.309)	(0.656)
$NCLB_{2008} \times T_s$	0.175	0.261	-0.046	-1.706*
	(0.312)	(0.741)	(0.467)	(0.833)
Sample N		5,710		
Interaction effects on tested teachers				
$NCLB_{2004} \times T_s$	0.183	0.649	0.271	1.522*
	(0.305)	(0.410)	(0.303)	(0.649)
$NCLB_{2008} \times T_s$	0.194	0.119	-0.021	-1.610*
	(0.316)	(0.721)	(0.467)	(0.83)
$NCLB_{2004} \times T_s \times Tested$	-0.447	0.546	-0.556	-2.786*
	(0.428)	(0.641)	(0.503)	(1.123)
$NCLB_{2008} \times T_s \times Tested$	0.239	0.531	-0.421	-1.578
	(0.403)	(1.017)	(0.528)	(1.93)
Sample N		9,270		

Notes. † $p \leq 0.1$; * $p \leq 0.05$; ** $p \leq 0.01$

(a) This model includes main effects only. Controls include teacher's gender, race/ethnicity, age, being a new teacher who taught three years or less, held a master's degree, state certified, entered the profession through alternative pathways, union status, and taught tested subjects in tested grades. It also includes school-level covariates (e.g., small or large small, disadvantaged schools, urban or rural school, and district instructional expenditure per pupil) and state-level covariates (e.g., state median household income and state employment population ratio).
(b) State fixed effects are included.
(c) Standard errors are clustered at state level and estimates are included in parentheses.
(d) Each model was estimated applying teacher follow-up final weights to make the results nationally representative and adjust for nonresponse bias.
(e) All sample sizes are rounded to the nearest ten due to the restricted-use data agreement with NCES.

Table 6. NCLB effects on disadvantaged schools

	Voluntary Mobility	Involuntary Mobility	Voluntary Attrition	Involuntary Attrition
Disadvantaged schools				
$NCLB_{2004} \times T_s$	0.25	1.548*	0.394	0.833
	(0.293)	(0.635)	(0.4)	(1.076)
$NCLB_{2008} \times T_s$	0.103	1.306†	0.068	-2.732**
	(0.279)	(0.803)	(0.325)	(1.004)
Sample N		4,550		
Advantaged schools				
$NCLB_{2004} \times T_s$	-0.105	0.437	0.350	-0.018
	(0.283)	(0.508)	(1.014)	(0.342)
$NCLB_{2008} \times T_s$	0.395	-0.769	0.487	-0.493
	(0.368)	(0.751)	(0.576)	(0.516)
Sample N		4,720		
Interaction effects on disadvantaged schools				
$NCLB_{2004} \times T_s$	-0.059	0.482	-0.088	0.421
	(0.282)	(0.464)	(0.338)	(0.93)
$NCLB_{2008} \times T_s$	0.403	-0.746	-0.466	-0.745
	(0.361)	(0.725)	(0.492)	(1.119)
$NCLB_{2004} \times T_s \times$ *Disadvantaged*	0.197	0.801	0.396	0.026
Schools	(0.345)	(0.623)	(0.563)	(1.514)
$NCLB_{2008} \times T_s \times$ *Disadvantaged*	-0.384	1.763†	0.404	-2.59†
Schools	(0.431)	(1.042)	(0.587)	(1.574)
Sample N		9,270		

Notes. † $p \leq 0.1$; * $p \leq 0.05$; ** $p \leq 0.01$

(a) This model includes main effects only. Controls include teacher's gender, race/ethnicity, age, being a new teacher who taught three years or less, held a master's degree, state certified, entered the profession through alternative pathways, union status, and taught tested subjects in tested grades. It also includes school-level covariates (e.g., small or large small, disadvantaged schools, urban or rural school, and district instructional expenditure per pupil) and state-level covariates (e.g., state median household income and state employment population ratio).
(b) State fixed effects are included.
(c) Standard errors are clustered at state level and estimates are included in parentheses.
(d) Each model was estimated applying teacher follow-up final weights to make the results nationally representative and adjust for nonresponse bias.
(e) All sample sizes are rounded to the nearest ten due to the restricted-use data agreement with NCES.

Appendix A. Coding Procedures for Treatment States

As shown in Table 1, the coding differentiates between states with state accountability systems prior to 2002 that resembled NCLB, and states that did not. We used a "0" to indicate that the state had the accountability system at the state level prior to 2002, and a "1" to signify that the state did not have such accountability system prior to 2002, so that an NCLB effect on schools and teachers after the implementation of NCLB was more likely to be observed. We began the coding procedure by evaluating a variety sources about pre-NCLB state accountability systems, and continued by referencing previous work by Hanushek and Raymond (2005), Dee and Jacob (2011), Carnoy and Loeb (2002), and Amrien and Berliner (2002). Both Hanushek and Raymond (2005) and Dee and Jacob (2011) code whether and when states had started consequential accountability prior to 2002. Because these two papers' coding is most consistent with the rationale of identifying NCLB treatment effects utilized in our study, we chose to compare our coding with that in these two papers as an illustration.

These papers differed in four states: Connecticut (CT), New Mexico (NM), North Carolina (NC) and Tennessee (TN). For CT, the two papers only disagreed on the year of accountability in the state, but both had years prior to NCLB. As a result, CT was coded as having an accountability system prior to the law. For NM, however, Dee and Jacob coded 1998 as the year of starting the state accountability system, while Hanushek and Raymond identified as 2003. Our review confirmed the 1998 date because there was a possibility of state sanctions, so we coded NM as "0". For NC, these two papers differed only in the year of accountability, but both 1996 and 1993 were well before NCLB. NC was also coded as "0" in our analysis. Similarly, Dee and Jacob did not agree with the year of accountability that Hanushek and Raymond listed for TN. While TN did start school performance reporting in 1996, they did not attach school-level sanctions until 2000. We used 2000 as the year for consequential accountability in TN, and coded it as a comparison state.

Two other states were coded differently in various papers: Kansas (KS) and Virginia (VA). For example, Hanushek and Raymond suggested that KS in 1995 only had a report card policy, rather than consequences attached to performance. However, as Dee and Jacob noted, KS did have, "an accreditation process that rated schools and could culminate in several possible sanctions for low-performing schools (i.e. closure)" (Dee and Jacob, 2011, appendix B). We verified this statement and coded KS as having pre-NCLB accountability. VA is unique, as it identified low-performing schools in 1998, but did not tie the loss of accreditation to the possibility of school closure (or other school sanctions). As a result, we chose to place VA in the treatment state group.

We also paid special attention to Alaska (AK), Arkansas (AR), and Illinois (IL). AK implemented a statewide accountability system in the 2001-02 school year. Since this came before NCLB was passed, we coded AK a comparison state. In terms of AR, both previous papers listed it as having an accountability system. However, Arkansas held districts accountable for performance, not individual schools. We determined that this would still possibly influence teacher employment decisions, and thus coded it as a comparison state. Dee and Jacob included IL as a comparison state, while Hanushek and Raymond did not. We chose to follow Dee and Jacob's coding, because IL had an early academic warning list for failing to meet the passing score on state tests for two consecutive years, or having a significant downward trend in test scores for three years. To us, this resembled NCLB enough to code IL as a "0".