Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students

By Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff*

Low-achieving students often are taught by the least-qualified teachers. These disparities begin when teachers take their first jobs, and in urban areas they are worsened by teachers' subsequent decisions to transfer and quit. Such quits and transfers increase disparities in at least two ways. First, more qualified teachers are substantially more likely to leave schools having the lowest-achieving students. For example, of the new teachers hired in New York City's lowest-achieving schools in 1996–1998, 28 percent scored in the lowest quartile on the general-knowledge certification exam.[1] Of those remaining in the same schools five years later, 44 percent had scores in the lowest quartile. In contrast, 22 percent of the new teachers in the higher-achieving schools were in the lowest quartile, which only increased to 24 percent for those remaining after five years.[2] Second, the generally high teacher turnover in lower-performing schools disadvantage students in those schools since the effectiveness of teachers increases over the first few years of their careers. Twenty-seven percent of first-year teachers in New York City's lower-performing schools do not return the following year, compared to 15 percent in the quartile of schools having the relatively highest student achievement.

* Boyd, Lankford, and Wyckoff: University at Albany, Albany, NY 12222; Loeb: Stanford University, Stanford, CA 94305. We are grateful to Richard Murnane for comments and to the Smith Richardson Foundation and the U.S. Department of Education for financial support. We also appreciate assistance with data from the New York State Education Department. None of these organizations necessarily supports the views expressed in this paper. Any errors are attributable to the authors.

[1] School performance is based on the 4th-grade English Language Arts exam.

[2] The 28 versus 22 percent comparison for entering (mostly certified) teachers understates the actual difference in the qualifications of new teachers across schools. For example, approximately half of the new teachers in the lowest-achieving schools were not certified, compared to 20 percent in the higher-achieving schools.

In this paper we examine New York City elementary school teachers' decisions to stay in the same school, transfer to another school in the district, transfer to another district, or leave teaching in New York State during the first five years of their careers. Our model allows us to go beyond past research in three important ways: examining how transfer and quit behavior is influenced by (i) interactions between teacher qualifications and school-level student achievement, (ii) unobserved heterogeneity in teachers' responses to school-level student attributes, and (iii) the distance from new teachers' prior homes to their initial job.

Many factors influence teacher transfers and quits. Teachers respond positively to increased salary, although the variation in salary across districts often is not large enough to strongly affect teacher sorting.[3] Non-pecuniary job characteristics such as class size, preparation time, facilities, student characteristics, and school leadership also can affect teacher decisions, and differences in these characteristics can be great across schools, especially in large urban areas (see Eric Hanushek et al., 1999; Loeb et al., 2005). In addition, teachers prefer schools and districts similar and geographically close to those they attended in high school (Boyd et al., 2004, 2005).

Several recent state-specific studies have considered student characteristics. Benjamin Scafidi et al. (2003) use a competing-risk model and find that Georgia elementary teachers move from schools with higher proportions of minority students and from low-performing schools, but the latter is explained by teacher preferences for fewer minority students. Hanushek et al. (2004), using a similar model and Texas data, find that teachers prefer higher-achieving students even after controlling for student racial composition. Both studies interact teacher race

[3] Recent studies include Peter Dolton and Wilbert van der Klaauw (1999) and Todd Stinebrickner (2000).

with student racial composition, allowing teachers of different races to have different preferences for student race.

We build on this work, examining teacher transfers and quits in a discrete-time competing-risk model with the extensions mentioned above. To our knowledge, this is the first study to look at the role that distance plays in teacher turnover and to employ a general mixed-logit framework that allows for very general unobserved heterogeneity in models of job transition.[4]

I. Data and Descriptive Analysis

Our data include information on the attributes of every teacher employed in a New York State public school at any time from 1995–1996 through 2003–2004, the environments in which they make career decisions (e.g., school attributes), and their locations at various points during their lives. Given our focus on student characteristics rather than salaries, it is advantageous to use data for a single large district. All New York City teachers are under the same salary schedule, avoiding the potential endogeneity of salary with teacher characteristics and working conditions. In addition, by modeling a single district we reduce the problems caused by teachers in different areas having different alternative job opportunities.

We examine transfers and quits by New York City teachers who started teaching in 1995–1996 through 2001–2002.[5] We limit the analysis to certified teachers with 1–5 years of experience, because attrition rates in these early years are highest, and New York City is no longer allowed to hire uncertified or "temporarily licensed" teachers. Furthermore, exit-survey data suggest that transfer and quit behavior of uncertified teachers is quite different from that of certified teachers.

Transfers and quits by New York City teachers are more common among those who teach low-performing and nonwhite students, have higher qualifications, are male, and lived outside New York City prior to teaching. The interaction between teacher qualifications and student achievement is particularly striking: highly qualified teachers are more likely to quit or transfer than less-qualified teachers, especially if they teach in low-achieving schools. For example, 20 percent of new teachers in the top quartile on the general-knowledge certification exam leave high-achieving schools after one year, but 34 percent of those in low-achieving schools leave after one year. By contrast, 14 percent of bottom-quartile teachers leave high-achieving schools after one year, and 17 percent leave low-achieving schools.

Finally, teachers who were living in New York City when they applied for certification, typically prior to their first teaching job offer, are far less likely to transfer to another school district and are less likely to quit than are their colleagues who were living elsewhere: fewer than 1 percent of New York City resident teachers transferred to another district after their first year, compared with almost 10 percent of nonresidents. In addition, distance to work matters: 7 percent of first-year teachers whose pre-certification address was within two miles of the school transfer to other district schools, compared with 11 percent of those six or more miles away.

The percentages of new teachers who are from New York City are lowest in low-performing schools, and their distances to the schools are greater. The most highly qualified new teachers in each student-performance group are less likely to be from New York City, and the new teachers who are residents have somewhat longer distances to their schools. Some of the higher turnover in low-performing schools may reflect teachers' preferences for proximity to home instead of preferences for higher-achieving students.

III. Methods

We model whether an individual initially teaching in a particular school (i) remains in that school, (ii) transfers to another school in the same district, (iii) transfers to another district, or (iv) leaves the New York City public school system, either exiting the labor force or taking alternative employment. Our discrete-time competing-risk model reflects both the annual nature of the data and the fact that most job

[4] Torberg Falch and Marte Ronning (2004) include a binary variable for Norwegian teachers born in the labor market in which they work.

[5] We define a "quit" as someone who has been missing from the data for two consecutive years.

transitions occur at the end of the school year.[6] To simplify the analysis, we focus on the first job change and track individuals up to five years.

If free to choose among all alternatives, rational choice implies that an individual will choose that which yields the highest satisfaction. Let $U_{it}^k = V_{it}^k + \varepsilon_{it}^k$ represent the utility of individual i in year t associated with alternative k where V_{it}^k is a linear function of the alternative's attributes. Assuming that the ε_{it}^k are drawn from the Gumbell distribution yields the multinomial-logit competing-risks model in (1). The specification in (2) includes attributes of both the teacher, z_{it}, and the school, x_{it}:

$$(1) \qquad p_{it}^{k'} = p^{k'}(x_{it}, z_{it} | \boldsymbol{\delta}_i, \boldsymbol{\eta}_i) = \frac{e^{\tilde{V}_{it}^{k'}}}{1 + \sum_{k=2}^{4} e^{\tilde{V}_{it}^{k}}} \qquad k' = 2, 3, 4$$

$$(2) \qquad \tilde{V}_{it}^k = V_{it}^k - V_{it}^1 = \beta_0^k + \beta_1^k z_{it} + \beta_2^k x_{it} + z_{it}' \theta^k x_{it} + \tau_t^k + \delta_i^k x_{it} + \eta_i^k.$$

The term $z_{it}' \theta^k x_{it}$ allows for observed heterogeneity in how the effects of various school attributes vary with individuals' own characteristics. In addition, the model allows for unobserved heterogeneity in the effects of school attributes via the random effect δ_i^k. The additive individual effect η_i^k allows for other unobserved heterogeneity. Others have estimated discrete-time hazard models that correspond to the case where all the teacher effects are zero. Here the elements of $\boldsymbol{\delta}_i$ and $\boldsymbol{\eta}_i$ are assumed to be independent normal random variables having zero means and standard deviations, with joint distribution, $f(\cdot)$. This formulation implies the likelihood of observing any particular job spell. For example, the expression in (3) is the probability that a new teacher remains in the same school three years and then transfers to another district:

$$(3) \qquad \int p_{i,t^0}^1 p_{i,t^0+1}^1 p_{i,t^0+2}^3 f(\cdot)\, d\boldsymbol{\delta}_i\, d\boldsymbol{\eta}_i.$$

[6] The data do not distinguish between transitions made during versus at the end of the school year.

Because our model falls within the mixed-logit framework discussed by David Revelt and Kenneth Train (1998), parameter estimates were obtained using their simulated maximum-likelihood estimation package.

Even though we introduced our model only referring to a teacher's choice of whether to transfer, the transition probabilities must be viewed as reduced-form specifications, as they will also depend on factors affecting the opportunities to transfer. For example, over the period studied, enrollment growth and increasing retirements yielded a tighter teacher labor market and more opportunities to transfer. Such trends will be reflected in the estimates of the year fixed effects, τ_t^k. Other estimated parameters also may reflect demand-side factors (e.g., teacher qualifications may influence opportunities to transfer). Thus, parameter estimates must be interpreted with caution. Even so, the model can provide some insights into behavior. For example, if a job candidate can remain in his current school and if other schools base their hiring decisions on the candidate's attributes and not those of his current school, then the estimated effects of current-school attributes would reflect only the teacher's preferences, not those of school officials.

Because each transition is associated with a composite alternative that aggregates "elemental alternatives" (e.g., individual schools), $\tilde{V}_{it}^k \equiv V_{it}^k - V_{it}^1$ in (1) reflects differences between the attributes of the initial school and the kth composite, possibly measured by summary measures. Prior empirical analyses of teacher transfers and quits have included measures characterizing current jobs but not alternative jobs. Yet a particular school might be an attractive alternative in one district but unattractive in another setting. If attributes of a school are correlated with omitted attributes of alternative schools, the estimated effects of the school's own attributes could be biased substantially. Because we are focusing on a single large district, individuals face a common set of schools in each composite alternative. Thus, the alternative-specific constant and year fixed effects will absorb the effects of variables characterizing the alternative schools.

IV. Results

First we estimate a baseline model similar to previous research, which includes student char-

acteristics, teacher characteristics, and interactions between teacher race and the racial composition of students. In addition, it includes whether the teacher was a resident of New York City when applying for certification and, for residents, distance from their home to their school.

Consistent with earlier studies, the level of student achievement and the racial composition of the student body are important determinants of the transition probabilities. In addition, interactions of teacher race with the racial composition of students are evident, especially for the probability of transferring within New York City. Simulated probabilities for this baseline model show that teachers in schools where student achievement is low are more likely to leave that school, most frequently transferring to a different school within the city, but the effects are not large.[7] White and Hispanic teachers are much more likely to leave schools as the proportion of white students falls and the proportion of black students increases. For black teachers there is little relationship between student-body racial composition and retention.[8]

The large difference in how teachers of different races react to student race is a good example of why it can be important to investigate heterogeneity. To understand this better we introduce two potential sources of observed heterogeneity: the interaction of teacher qualifications (teacher certification scores and experience) with school-level student achievement, and the distance between the residence of a teacher immediately prior to beginning her career and the location of her first teaching assignment. We also model unobserved heterogeneity in the effects of (a) student-body racial composition and (b) the proportion of a school's students who fail the English Language Arts (ELA) 4th-grade exam, both captured by $\boldsymbol{\delta}_i$. We also include generic unobserved heterogeneity, $\boldsymbol{\eta}_i$. The simulation results for new teachers in this extended model are in Table 1A. (Coefficient estimates for both models and simulation results for teachers having prior experience are available from the authors upon request.)

TABLE 1—PROBABILITIES OF RETENTION, TRANSFER, AND QUITS FOR SELECTED TEACHER AND STUDENT ATTRIBUTES BASED ON MODEL-2 ESTIMATED COEFFICIENTS, FIRST-YEAR TEACHERS

Attribute and statistic	Remain in same school	Transfer within NYC	Transfer outside of NYC	Leave NYC schools
A)				
Percentage of Students Failing ELA for Top 25 percent Scoring Teachers				
−1 SD (40 percent)	0.807	0.070	0.018	0.105
Mean (60 percent)	0.790	0.080	0.018	0.112
+1 SD (80 percent)	0.768	0.091	0.019	0.122
Percentage of Students Failing ELA for Middle 50 percent Scoring Teachers				
−1 SD (40 percent)	0.852	0.071	0.014	0.063
Mean (60 percent)	0.826	0.079	0.017	0.078
+1 SD (80 percent)	0.795	0.087	0.021	0.097
Percentage of Students Failing ELA for Bottom 25 percent Scoring Teachers				
−1 SD (40 percent)	0.844	0.073	0.013	0.070
−1 (60 percent)	0.853	0.071	0.013	0.063
+1 SD (80 percent)	0.858	0.068	0.012	0.062
B)				
Percentage of Students Failing ELA, Allowing for Unobserved Heterogeneity, Middle 50 percent Scoring Teachers				
Minus 1 standard deviation in δ^4 for percent failing:				
−1 SD (40 percent)	0.890	0.074	0.014	0.022
Mean (60 percent)	0.884	0.084	0.018	0.014
+1 SD (80 percent)	0.874	0.095	0.023	0.009
Average unobserved heterogeneity for percent failing:				
−1 SD (40 percent)	0.865	0.072	0.014	0.050
Mean (60 percent)	0.854	0.082	0.018	0.047
+1 SD (80 percent)	0.842	0.092	0.022	0.045
Plus 1 standard deviation in δ^4 for percent failing:				
−1 SD (40 percent)	0.833	0.065	0.012	0.090
Mean (60 percent)	0.791	0.071	0.014	0.124
+1 SD (80 percent)	0.766	0.073	0.016	0.145
C)				
Distance, New York City residents				
0 miles	0.809	0.071	0.018	0.102
3 miles	0.794	0.078	0.018	0.110
6 miles	0.778	0.085	0.019	0.117
10 miles	0.756	0.096	0.020	0.128
Reside outside NYC	0.682	0.063	0.125	0.129
D)				
School Composites, Deciles Ranking by Student ELA Failure Rate				
Higher achieving (2nd)	0.851	0.062	0.015	0.073
Middle case (5th decile)	0.792	0.074	0.019	0.115
Lower achieving (9th)	0.741	0.097	0.021	0.141

[7] Unless otherwise specified, the reference teacher is white with one year of experience and a master's degree, in the top quartile of the general-knowledge certification exam, and, prior to beginning to teach, lived in New York City 3.8 miles from the school where she currently teaches. School attributes are held at their means.

[8] The estimated models also include the percentages of students receiving free lunch and having limited English proficiency, school enrollment, and dummy variables indicating a teacher's sex, years of experience, grade, subject taught, and year.

When teacher performance on the general knowledge certification exam is interacted with the proportion of students in a school who fail the ELA exam, a striking pattern emerges. The top 75 percent of teachers react much more strongly to low-achieving students than do the bottom 25 percent. Retention of first-year teachers in the top quartile falls by 4 percentage points when the portion of students failing the ELA exam varies from one standard deviation below the mean to one standard deviation above, and falls by about 6 percentage points for those in the middle 50 percent. By contrast, retention of teachers in the lowest quartile increases slightly.

Even after accounting for differences due to prior teaching experience, certification exam performance, and their interaction, there is substantial unobserved heterogeneity in how the school-level student achievement affects transfers and quits; the estimated standard deviation for the element of $\boldsymbol{\eta}_i^4$ corresponding to student achievement is statistically significant (robust $t = 2.44$) and large (2.12). To investigate this unobserved heterogeneity, we simulate estimates of the transition probabilities where the integration in expressions such as (3) is with respect to all the other random effects. Table 1B shows results for one-standard-deviation differences in this random effect. We find large differences across teachers: some appear not to consider student performance in their career decisions, while others are far more likely to leave their school when the proportion of low-performing students is higher. For example, those teachers one standard deviation higher in unobserved heterogeneity are approximately 7 percentage points more likely to leave a school with one standard deviation more low-performing students than a school with one standard deviation fewer low-performing students relative to the mean.

Race and ethnicity of teachers and their students are also important determinants of teacher retention, with the results very similar to those of the baseline model. Once we account for the observed heterogeneity through teacher–student race interactions, the estimated standard deviations for student race and ethnicity are small in magnitude and statistically insignificant.

School attributes are typically correlated, which compounds the difficulty of staffing schools where academic performance is low. Simulations for simultaneously altering all the observable attributes of schools to reflect typical values for schools at various points in the student academic achievement distribution are shown in the last panel (d) of Table 1. These changes in school attributes imply substantial changes in teacher retention. For both new and more experienced teachers, moving from schools at the 20th percentile to those at the 80th percentile reduces retention by over 10 percentage points. For new teachers, quits double, and within-district transfers increase by about 50 percent. For more-experienced teachers, most of the reduced retention is caused by increasing transfers among schools in New York City.

The location of a teacher immediately prior to beginning her teaching career is also an important determinant of retention. New York City residents who lived further from their schools are more likely to leave (Table 1C). There is an even larger difference reflecting whether or not an individual was a New York City resident. Nonresidents are five times more likely to transfer to positions outside New York City, both after their first year of teaching and in subsequent years. We estimate that a teacher who lived outside New York City prior to beginning his or her teaching career only has a 20 percent probability of remaining in the same school beyond five years, and a probability of transferring to a school outside New York City that exceeds 30 percent. By contrast, a teacher who lived in New York City about three miles from school is twice as likely to remain in the same school.

The differences related to geography are as large as any of the other factors considered. These findings are consistent with our earlier findings that teachers much prefer to locate in areas closer to home and in schools similar to those where they went to high school. (Boyd et al., 2004, 2005). The causality associated with the relationship between location and transfer is unclear. Nonetheless, geography provides a good indicator of the likelihood of transfer.

V. Conclusion

This paper makes three contributions to our understanding of why low-achieving students often are taught by the least-qualified teachers. First, interactions between teacher qualifications and

student achievement are important. Teachers, especially highly qualified teachers, are more likely to transfer or quit when teaching lower-achieving students, even after accounting for student and teacher race. Second, unobserved heterogeneity is important: although the average teacher apparently prefers to avoid schools with low-performing students, many teachers are unaffected while others are strongly affected. Third, the geography of prior home and job is important. Teachers who lived farther from their school prior to beginning their job are more likely to quit or transfer. This phenomenon has substantive implications as urban districts are net importers of teachers from surrounding suburbs. For example, in the period studied, 34 percent of all certified teachers newly hired in New York City and 38 percent of those scoring in the highest quartile on the general-knowledge certification exam resided outside New York City at the time they sought certification. Distance helps explain the high quit rates of teachers in urban schools.

The policy implications of our findings are not straightforward. While teachers are leaving schools with low-performing students, we do not know whether this is due to the students or to correlates of student composition. In addition, we cannot tell whether school leadership strategies, hiring practices, pre-service preparation, or mentoring and induction could change the evident patterns. These factors also could explain why there is meaningful unobserved heterogeneity in teachers' responses to observed student performance. Important questions remain as to whether this heterogeneity reflects teachers having different responses to similar school environments or schools having similar student performance being in other ways quite different. For this we need to supplement this data with more detailed data on schools and teachers.

REFERENCES

Boyd, Donald; Lankford, Hamilton; Loeb, Susanna and Wyckoff, James. "Analyzing the Determinants of the Matching of Public School Teachers to Jobs." Working paper, University of Albany, 2004.

Boyd, Donald; Lankford, Hamilton; Loeb, Susanna and Wyckoff, James. "The Draw of Home: How Teachers' Preferences for Proximity Disadvantage Urban Schools." *Journal of Policy Analysis and Management*, 2005, *24*(1), pp. 1134–32.

Dolton, Peter J. and van der Klaauw, Wilbert. "The Turnover of Teachers: A Competing Risks Explanation." *Review of Economics and Statistics*, 1999, *81*(3), pp. 543–52.

Falch, Torberg and Rønning, Marte. "The Influence of Student Achievement on Teacher Turnover." Working paper, Norwegian University of Science and Technology, Trondheim, 2004.

Hanushek, Eric A.; Kain, John F. and Rivkin, Steven G. "Do Higher Salaries Buy Better Teachers?" National Bureau of Economic Research (Cambridge, MA) Working Paper No. 7082, 1999.

Hanushek, Eric A.; Kain, John F. and Rivkin, Steven G. "Why Public Schools Lose Teachers." *Journal of Human Resources*, 2004, *39*(2), pp. 326–54.

Loeb, Susanna; Darling-Hammond, Linda and Luczak, John. "How Teaching Conditions Predict Teacher Turnover in California Schools." *Peabody Journal of Education*, 2005 (forthcoming).

Revelt, David and Train, Kenneth. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics*, 1998, *80*(4), pp. 647–57.

Scafidi, Benjamin; Stinebrickner, Todd and Sjoquist, David L. "The Relationship Between School Characteristics and Teacher Mobility." Working paper, Georgia State University, 2003.

Stinebrickner, Todd. "An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Quit." *Journal of Human Resources*, 2002, *37*(1), pp. 192–216.