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

Donald Boyd
Milne 300
135 Western Ave
University at Albany
Albany, NY 12222
boydd@rockinst.org
518-443-5284
518-443-5274 (fax)

Hamilton Lankford
Milne 300
135 Western Ave
University at Albany
Albany, NY 12222
hamp@albany.edu
518-442-2621
518-442-3898 (fax)

Susanna Loeb
224 CERAS
520 Galvez Mall
Stanford University
Stanford, CA 94305
sloeb@stanford.edu
650-725-4262
650-723-7578 (fax)

James Wyckoff
Milne 300
135 Western Ave
University at Albany
Albany, NY 12222
wyckoff@albany.edu
518-442-5269
518-442-3898 (fax)

January 10, 2005

* A longer version of this paper is available at www.teacherpolicyresearch.org
Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students*

Low-achieving students often are taught by the least-qualified teachers. These disparities begin when teachers take their first jobs, and in urban areas 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 1996-98, \(^1\) 28 percent scored in the lowest quartile on the general knowledge certification exam. 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 teacher 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.

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 (1) interactions between teacher qualifications and school-level student achievement, (2) unobserved heterogeneity in teachers’ responses to school-level student attributes, and (3) 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. 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. In addition, teachers prefer schools and districts similar and geographically close to those they attended in high school (Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff, 2004 and 2005).

Several recent state-specific studies have considered student characteristics. Benjamin Scafidi, Todd Stinebrickner, and David Sjoquist (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. Eric Hanushek, John Kain and Steven Rivkin (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 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.

II. Data and descriptive analysis

Our data includes information on the attributes of every teacher employed in a New York State public school at any time from 1995-96 through 2003-04, 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 1995-96 through 2001-2002. We limit the analysis to certified teachers with one to five 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 one percent of NYC resident teachers transferred to another district after their first year, compared with almost ten percent of non-residents. In addition, distance to work matters: seven percent of first-year teachers whose pre-certification address was within two miles of 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 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 (1) remains in that school; (2) transfers to another school in the same district; (3) transfers to another district, or (4) leaves the NYS public 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 transitions occur at the end of the school year.

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 an individual will choose that which yields the highest satisfaction. Let \( U_{it}^k = V_{it}^k + \epsilon_{it}^k \) 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 \( \epsilon_{it}^k \) 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, \( x_{it} \).

\[
(1) \quad p_{it}^{k'} = p^{k'}(x_{it}, z_{it} \mid \delta_i, \eta_i) = \frac{\tilde{e}_{it}^{k'}}{1 + \sum_{k=2}^{4} \tilde{e}_{it}^{k'}}; \quad k' = 2, 3, 4
\]

\[
(2) \quad \tilde{V}_{it}^{k} = V_{it}^{k} - V_{it}^{1} = \beta_{0k} + \beta_{1k} z_{it} + \beta_{2k} x_{it} + z_{it}' \theta_{k} x_{it} + \tau_{it}^{k} + \delta_{it}^{k} x_{it} + \eta_{it}^{k}
\]
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 \( \delta^k \). The additive individual effect \( \eta^k \) 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 \( \delta \) and \( \eta \) 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. 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.

\[
(3) \quad \int p_1^{i,1} p_1^{i,2} p_1^{i,3} f(\cdot) d\delta d\eta
\]

Even though we introduced our model only referring to a teacher’s choice 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, \( \tau^k \). 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), $r_{it}^k = r_{it}^k - r_{it}^1$ 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 characteristics, 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. White and Hispanic teachers are much more likely to leave schools as the proportion of white students falls and the proportion black students increases. For black teachers there is little relationship between student-body racial composition and retention.
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 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 ELA 4th grade exam, both captured by $\delta_i$. We also include generic unobserved heterogeneity, $\eta_i$. The simulation results for new teachers in this extended model are in Table 1. (Coefficient estimates for both models and simulation results for teachers having prior experience are available upon request.)

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 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 $\eta_i^4$ 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 2 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 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. There is an even larger difference reflecting whether or not an individual was a New York City resident. Non-residents 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 their teaching careers 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 3 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, Lankford, Loeb & Wyckoff, 2004 and 2005). Even when a greater distance reflects a teacher’s plans to move in the future, it nonetheless, is 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: while 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


<table>
<thead>
<tr>
<th>Simulation</th>
<th>Remain in same school</th>
<th>Transfer within NYC</th>
<th>Transfer outside of NYC</th>
<th>Leave NYS schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Students Failing for Top 25% Scoring Teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.807</td>
<td>0.070</td>
<td>0.018</td>
<td>0.105</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.790</td>
<td>0.080</td>
<td>0.018</td>
<td>0.112</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.768</td>
<td>0.091</td>
<td>0.019</td>
<td>0.122</td>
</tr>
<tr>
<td>% Students Failing for Middle 50% Scoring Teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.852</td>
<td>0.071</td>
<td>0.014</td>
<td>0.063</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.826</td>
<td>0.079</td>
<td>0.017</td>
<td>0.078</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.795</td>
<td>0.087</td>
<td>0.021</td>
<td>0.097</td>
</tr>
<tr>
<td>% Students Failing for Bottom 25% Scoring Teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.844</td>
<td>0.073</td>
<td>0.013</td>
<td>0.070</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.853</td>
<td>0.071</td>
<td>0.013</td>
<td>0.063</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.858</td>
<td>0.068</td>
<td>0.012</td>
<td>0.062</td>
</tr>
<tr>
<td>% Students Failing Allowing for Unobserved Heterogeneity (middle 50% scoring teachers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MINUS 1 standard deviation in $\delta^4$ for percent failing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.890</td>
<td>0.074</td>
<td>0.014</td>
<td>0.022</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.884</td>
<td>0.084</td>
<td>0.018</td>
<td>0.014</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.874</td>
<td>0.095</td>
<td>0.023</td>
<td>0.009</td>
</tr>
<tr>
<td>Average unobserved heterogeneity for percent failing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.865</td>
<td>0.072</td>
<td>0.014</td>
<td>0.050</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.854</td>
<td>0.082</td>
<td>0.018</td>
<td>0.047</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.842</td>
<td>0.092</td>
<td>0.022</td>
<td>0.045</td>
</tr>
<tr>
<td>PLUS 1 standard deviation in $\delta^4$ for percent failing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1 std dev (40%)</td>
<td>0.833</td>
<td>0.065</td>
<td>0.012</td>
<td>0.090</td>
</tr>
<tr>
<td>Mean (60%)</td>
<td>0.791</td>
<td>0.071</td>
<td>0.014</td>
<td>0.124</td>
</tr>
<tr>
<td>+1 std dev (80%)</td>
<td>0.766</td>
<td>0.073</td>
<td>0.016</td>
<td>0.145</td>
</tr>
</tbody>
</table>

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 |

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 |
Endnotes

* Donald Boyd, University at Albany, Albany, NY 12222; Hamilton Lankford, University at Albany, Albany, NY 12222, Susanna Loeb, Stanford University, Stanford CA; James Wyckoff, University at Albany, Albany, NY, 12222. 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 support the views expressed in this paper. Any errors are attributable to the authors.

1 School performance is based on the fourth 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.

3 Recent studies include Peter Dolton and Wilbert van der Klaauw, 1999; Todd Stinebrickner, 2000.

4 See Eric Hanushek, John Kain and Steven Rivkin (1999), and Susanna Loeb, Linda Darling-Hammond and John Luczak (forthcoming).

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

6 We define a “quit” as someone who has been missing from the data for two consecutive years.

7 The data does not distinguish between transitions made during versus at the end of the school year.

8 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, who 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.

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