Available online at www.sciencedirect.com
SCIENCE DIRECT
Economics of Education Review 25 (2006) 13–27
Economics of Education Review
www.elsevier.com/locate/econedurev

Teacher quality and attrition

John M. Krieg*

Western Washington University, Bellingham, WA 98225-9074, USA

Received 14 May 2004; accepted 22 September 2004

Abstract

The argument that instructors with marketable skills are likely to exit the teaching profession leads many to believe that public schools are populated by teachers of mediocre talent. Yet, teachers with skills attractive to non-education employment may not be the best individuals in the classroom. A two-stage regression technique first estimates a teacher's impact on their students conditional upon prior academic achievement and then uses this quality measure to explain teacher attrition. This paper finds that higher-quality female teachers are less likely to leave the profession. Teacher quality does not impact attrition of male teachers.

JEL classification: I20; J44

Keywords: Educational economics; Human capital

> Conditions that undermine the power and effectiveness of the public school system need to be identified and promptly rectified... This includes, above all, creating a work environment that will continue to draw the bright, committed new teachers we need... But our track record over the past 40 years isn't very promising. Too many will quit permanently because they are fed up. Their ambition and self-respect will take them into business or other professions... They leave behind an increasing proportion of tired time-servers.
> Life, November 16, 1962.

1. Introduction

As evidenced by the above quote, a long-held belief is that public schools lose their best teachers to external job opportunities leaving behind below average instructors. These beliefs are reinforced by the difficulty schools face when attracting applicants for science, mathematics, and technical teaching positions. While teachers with skills attractive to other employers may be more likely to leave the profession, it is not clear that these are the "best" teachers. As a matter of fact, little research investigates the role of teacher quality on teacher attrition.

Recent studies reinforce the belief that teacher quality is an essential component of student success. Using student-level data, Hanushek, Kain, and Rivkin (1998) report that at least 7.5% of the total variation in student achievement is explained by teacher fixed effects. These teacher impacts are estimated to be larger than the effects of overall school organization, leadership, and financial conditions. Because of the obvious importance of teachers, school districts have pursued a wide array of strategies hoping to increase quality. The most common policy is to increase wages with the intent of attracting

*Tel.: +1 360 650 7405; fax: +1 360 650 4844.
E-mail address: john.krieg@wwu.edu (J.M. Krieg).

0272-7757/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.econedurev.2004.09.004

better candidates. Other districts offer signing bonuses, student loan forgiveness, or housing assistance programs. A handful of districts have even incorporated wage structures that reward high-performing teachers.

All of these strategies derive from the belief that quality, as measured by observable characteristics such as academic qualifications or certification results, is correlated with student achievement. Yet few studies find such a connection, suggesting the likelihood that observed teacher characteristics are only weakly, or not at all, correlated with quality. This is especially perplexing given the probable positive bias in estimated coefficients caused by high-quality teachers selecting into school systems with better students.

The purpose of this work is to investigate the impact of quality on one part of teacher labor supply: the decision to leave the profession. The standard argument is that more able teachers are attractive to outside employers and hence likely to leave schools. Of course this is conditional on the fact that teacher quality is correlated with skills demanded by non-school employers, a possibility that is likely true at higher grades but perhaps not at lower grades where skills focusing on young children are valued by education but less so by the business community.

On the other hand, arguments can be made for why more able instructors remain in education. A teacher who performs well in the classroom receives the respect of their colleagues, recognition from administrators, and possibly better future choices of students, buildings, districts, and teaching-related opportunities. Further, highly skilled individuals entering teaching typically value non-pecuniary rewards of teaching over the rewards of non-academic jobs. Attracting these individuals to non-academic opportunities likely requires greater compensation and hence a lower likelihood of non-academic job offers made to these individuals. Finally, it is commonplace for administrators to be promoted from the ranks of teachers and proficient instructors potentially increase their probability of promotion and are less likely to exit teaching.

Discussing the impact of teacher quality on exit decisions begs the question of quality measurement. The econometric model used in this paper is a two-stage regression technique, which first develops teacher quality measures by estimating the average gains on standardized tests made by students during the year spent with their teacher. A second-stage regression then employs these quality estimates as independent variables in logit models explaining attrition decisions. This test-based technique offers the benefit of directly measuring teacher impacts on students rather than indirectly measuring quality through observable teacher measures at the time of hire. This paper presents evidence that female instructors performing better in the classroom are less likely to exit the profession. Unlike women, no relationship between quality and attrition is found for men, suggesting different labor supply functions between genders.

2. Econometric strategy and estimation

2.1. Strategy

Recent empirical work increasingly uses estimated variables and coefficients from a primary regression as independent variables in a secondary regression. This two-stage process is discussed by Amemiya (1978), and in a classic example of this technique, Heckman (1979) determines the correct asymptotic covariance matrix of a two-stage estimator that accounts for sample selection bias. In a technique closely related to that used in this paper, Stoddard (forthcoming) first estimates area amenities through state-level fixed effects in a wage premium equation for non-teachers, and then uses these estimates in regressions explaining teacher wages.

The two-stage procedure used here is similar to that employed by Rockoff (2004). In his work, Rockoff measures teacher ability by explaining student achievement on standardized tests with teacher fixed effects. The teacher fixed effects are interpreted as direct measures of teacher value added or what this paper terms as quality. Rockoff finds that a teacher one standard deviation above average has students who perform between one- and two-tenths of a standard deviation above reading and mathematics standardized test averages. This paper extends Rockoff's technique by first estimating a teacher's impact on their students' standardized test scores and then employs the estimated impacts as independent variables in a logit model of teacher attrition. Specifically, the first stage investigates a panel of students within a teacher's classroom. This model is

$$A_{ij} = \gamma A_{i,j-1} + \boldsymbol{\beta}\mathbf{X}_{ij} + \boldsymbol{\delta}\mathbf{T}_j + Q_j + \varepsilon_{ij}, \qquad (1)$$

where A_{ij} is the achievement on a standardized exam earned by student i of teacher j, $A_{i,j-1}$ is the student's test score earned under their previous teacher, $\mathbf{X}$ is a matrix of observable student, building, and district characteristics, $\mathbf{T}$ is a matrix of teacher characteristics, Q represents teacher-specific effects, $\boldsymbol{\delta}$ and $\boldsymbol{\beta}$ are vectors of regression coefficients, and ε the regression error term. The variables of interest in Eq. (1) are the Q's and $\boldsymbol{\delta}\mathbf{T}_j$'s, which together represent the impact of a teacher on a student's test scores conditional upon previous test scores as well as observable student, building, and district measures.

As teacher quality is the sum of observable teacher-specific characteristics and unobservable characteristics captured by Q, quality is estimated by

$$\widehat{\text{Quality}}_j = \hat{Q}_j + \hat{\boldsymbol{\delta}}\mathbf{T}_j. \qquad (2)$$

Thus, the proposed measure of teacher quality is the average student test gain attributable to their teacher conditional upon the other regressors in Eq. (1). Since $\widehat{\text{Quality}}$ is generated from a regression explaining student test scores based upon previous student performance, $\widehat{\text{Quality}}$ may be interpreted as a particular teacher's value added to their student's education.

In his work, Rockoff measures teacher quality by estimating Eq. (1) absent measurable teacher characteristics, **T**. This allows Rockoff to estimate Q using ordinary least-squares fixed effects. By introducing measurable teacher characteristics, which do not vary within a classroom, this paper estimates Q with random effects techniques. Later, in order to check for robustness across estimation techniques, this paper also reports results using Rockoff's methodology.

To identify if high-quality teachers are more likely to exit the profession, the Quality estimates are incorporated into a logit model with dependent variable equaling one if a teacher exits the profession. Specifically, this model takes the form

$$\Pr[Y_j = 1] = f(\lambda \mathbf{Z}_j + \phi\, \widehat{\text{Quality}}_j + \upsilon_j), \tag{3}$$

where $\mathbf{Z}$ is a matrix of variables influencing attrition, λ and ϕ are regression coefficients, υ is a standard error term, and $\widehat{\text{Quality}}$ is determined by Eq. (2). If high-quality teachers are more likely to exit, ϕ will be positive.

Before estimating the aforementioned equations, a number of econometric concerns require attention. First, since quality is an estimated variable in Eq. (3), even if quality is measured without bias, random measurement error likely exists. The implications of errors in variables in linear models are well understood: coefficient estimates of mis-measured variables are attenuated towards zero while coefficients of other included variables are biased usually in unknown directions (Greene, 2000). The magnitude of this bias is directly related to the variance of the measurement errors. DeVaro and Lacker (1995) demonstrate that errors in variables also attenuate the coefficients of mis-measured variables in logit models towards zero. Thus, if the teacher quality estimates are measured with error, then the estimate of ϕ will represent a lower bound on the magnitude of the effect of quality on exits.

A more insidious possibility is that the measures of teacher quality generated in Eq. (1) are biased. Considering the first stage equation, a source of bias occurs when variables explaining student test scores are omitted. If omitted variables are positively correlated with both teacher quality and student test scores, then the first-stage estimates of quality will be biased upwards. Classic examples of bias arise from correlations between teacher quality and the level of community or parental involvement, the reputation of a school district, and the effort put forward by students. For instance, if high quality instructors are more likely to teach in districts with high parental involvement, then omitting parental involvement biases the estimates of teacher quality upwards. The impact of using biased regressors in the second stage regression is unclear, but under plausible circumstances should attenuate ϕ towards zero. To understand this, consider a positive bias in quality estimates for high-ability teachers. If these teachers have a natural propensity to exit based upon their quality yet a researcher incorrectly overstates their quality, then the researcher will underestimate the impact of a unit of quality on the probability to exit. Of course if the bias occurs equally across all levels of quality, no bias in the second stage slope parameters will occur.

Scenarios other than mis-measured Quality that result in biasing ϕ can be constructed. Consider a building with a poor quality administrator whose leadership simultaneously causes students to underachieve and teachers to be more likely to exit. Based upon the underachieving students, a researcher employing Eq. (1) would identify that building's teachers as low quality. Since those teachers are also more likely to exit, a researcher would incorrectly find a negative impact of quality on attrition. Another possibility is that principals in some buildings match strong students with more able teachers and weaker students with weaker teachers. If all teachers are less likely to exit when given high ability students, then this sorting will lead to higher exit rates among low quality teachers caused not by their quality but rather by the sorting of students. In order to control for these types of scenarios, alternative specifications combine building dummy variables in the first-stage regressions with building-level random effects logit models. These results are presented after the initial regression logit estimates.

A final econometric concern regards the inference of ϕ. Typical standard errors in logit models are estimated based upon the regressors being correctly measured. As the quality variable is a generated one, the conditional logit standard errors are likely to be biased towards zero. In order to correct this, Gawande's (1997) technique of adjusting the standard errors of generated regressors is employed.[1]

[1]Gawande suggests creating an adjusted regressor according to:

$$\widehat{\text{Quality}}_i = \overline{\widehat{\text{Quality}}} + \hat{\sigma}_Q^2/\sigma_{u,i}^2\left(\widehat{\text{Quality}}_i - \overline{\widehat{\text{Quality}}}\right)$$

where $\sigma_{u,i}^2$ is the measurement error of the generated variable and $\hat{\sigma}_Q^2$ is an estimate for the sample variance of Q had it been measured without error. $\widehat{\text{Quality}}$ are then used in the second stage logit estimation.

2.2. Teacher quality estimation

Under the No Child Left Behind Act (NCLBA), each state is required to test student achievement. The Washington Assessment of Student Learning (WASL) is the state of Washington's diagnostic tool used for identifying failing schools and students under the NCLBA. The WASL is an open-ended exam given near the end of the academic year that covers four subjects: reading, writing, listening and mathematics.[2] The WASL is administered in grades 4, 7, and 10 and, under current legislation, students are required to pass the WASL in order to receive a high school diploma. Four variants of Eq. (1) are estimated: one each with the individual 4th grade student's WASL reading, writing, listening, and mathematics scores used as dependent variables. To make comparisons with other standardized tests easier, WASL scores are normalized so that the mean of the observations is zero with standard deviation of one.

Students in Washington are also required to complete the Iowa Test of Basic Skills (ITBS). The Iowa tests are standardized exams intended to identify a student's developmental level and to measure annual academic growth. The ITBS is given in Washington near the end of the student's 3rd and 6th grade years and covers four disciplines: reading, vocabulary, listening, and mathematics. For each variant of Eq. (1), the student's 3rd grade ITBS discipline score is matched to the corresponding WASL discipline score and employed as an independent variable.[3] Relative to Eq. (1), the $A_{i,j-1}$ are measured by ITBS performance and the $A_{i,j}$ are WASL results. Thus, Eq. (1) measures student performance at the end of their 4th grade year conditional upon their academic achievement at the end of the 3rd grade as well as other student, building, classroom and teacher characteristics.[4]

The data were provided by Washington's Office of the Superintendent of Public Instruction and consist of two files: a student file and a teacher file. The student file is a panel of all students who took both the 3rd grade ITBS in the 2000–2001 academic year and the 4th grade WASL in 2001–2002. The student file lists both the building and district in which the student was enrolled as well as his or her teacher's first and last name. Because children in higher grades typically receive core subject material from many teachers, this work analyzes only the cohort of 3rd grade students who took the ITBS and the subsequent WASL. The student file also contains a wealth of demographic and social data generated from student questionnaires associated with the WASL and ITBS.

A second data set, the teacher file, represents a complete panel of annual observations of Washington public school teachers between 1996 and 2004. The teacher data set contains the building and district in which the teacher works, teacher experience, education, salary, and demographic data as well as a unique teacher identification number that remains constant throughout the teacher's career. After using Eq. (1) to estimate the teacher quality variables, these variables were merged into the teacher file using the teacher's building and first, middle, and last names. Then, using the unique teacher identification number, each teacher was determined to either remain in the profession after the 2001–2002 academic year or to have exited the profession. Because leaves of absence are common in education, a teacher was determined to have exited the profession only if they did not teach in either of 2002–2003 or 2003–2004 academic year.[5] The end result is a cross section of teachers who administered the 2001–2002 WASL, which also contains the four subject-quality measures as well as a zero-one identifier indicating teacher exit. After excluding missing observations from both the student and teacher samples as well as students who could not be matched with their fourth grade teacher, 36,056 observations remain in the student file.[6] These students are matched with 2293 instructors in the teacher file.

---

[2]In 2004, the state of Washington decided to replace the listening test with a science test. Statewide results for this test are not yet available.

[3]The WASL writing score is matched to the ITBS vocabulary score.

[4]Because the WASL and ITBS are two different tests scored on completely different scales, I chose to use the ITBS as an explanatory variable on WASL performance rather than simply attributing the difference between the two tests to the student's fourth grade teacher. Like the WASL scores, all ITBS results are normalized to mean zero and standard deviation equal to 1.

[5]This paper does not follow teachers into either private schools nor public instruction in another state. Thus the number of teachers exiting the profession may overstate the amount leaving the profession because some exits from public education may go to private schools. Two years was used as the cutoff to determine teacher exit because most school districts in the state of Washington grant leaves of absence only up to 1 year in length.

[6]The sample of 36,056 students represents nearly half of the state's 78,610 fourth graders in the 2001–2002 school year. Roughly 32,000 students were excluded because of an inability to match with certainty the student's classroom teacher with a teacher in the personnel records. This inability to match is random in nature; classroom teachers were matched to personnel records based upon last names, first names, and buildings in which they worked. Because of misspelling of teacher names and names which could be attributed to more than one classroom teacher (e.g. Smith), a certain match was not made and both the student and teacher data were omitted. A further 8000 students were excluded from the sample because of missing student data as were all teachers who job shared, worked less than full-time, taught in split classrooms, or left mid-year. Because one reason for missing student data is due to student mobility, this paper may under count this type of student.

In order to accurately capture the gains made under their 4th grade teacher, Eq. (1) includes student's 3rd grade ITBS results. Other measures of student characteristics in Eq. (1) include the length of time the student has been enrolled in their building, the presence and use of a computer at home, the frequency of both reading for fun and of watching television, English use at home, migrant status, gender, and race. Measures specific to the student's building and classroom are also included. These consist of the student's class size, the average ITBS subject score for the student's classmates, the percent of a building's students on free or reduced lunch, and the percent of the building's students passing the previous year's WASL. Both the percent of previous WASL-passers (in the student's building) and the student's classmates' average ITBS subject score proxy for peer effects, which may influence student performance during their WASL year. Finally, 296 district dummy variables are included to capture heterogeneity in student performance across districts.

Each column of Table 1 presents selected coefficient estimates of Eq. (1) using the four different WASL subject scores as dependent variables.[7] Not surprisingly, students performing well on the matching 3rd grade ITBS subject test do well on their 4th grade WASL. For instance, a student scoring one standard deviation above the 3rd grade math ITBS mean is expected to place .713 standard deviations above the 4th grade math WASL mean. Students new to the school, those previously held back a grade, blacks, Hispanics, and males average lower scores across subjects. Likewise, students having computers at home (possibly a proxy for income) and those who frequently read for fun score higher across subjects. A student's environment is found to impact their test performance; those in buildings where previous students performed well are likely to perform better as are students whose classmates performed well on the ITBS. Observable teacher characteristics also impact student performance; students of male teachers and first year teachers average lower test scores. Finally, teachers with masters degrees perform no differently than those with bachelor's degrees, and experience impacts student test scores in non-linear manner; the marginal impact of experience diminishes as experience increases.

The purpose of the first stage regressions is to generate four different quality estimates based upon Eq. (2): one for each WASL subject. Histograms of the resulting quality estimates sorted by attrition status are presented in Fig. 1. The average teacher quality generated by Eq. (2) range from .010 (writing) to .047 (listening) with standard deviations ranging from .123 (listening) to .272 (mathematics). As the dependent variable in Eq. (1) is the standardized performance on the WASL conditional upon the ITBS, the interpretation of quality is measured as the average student gains on the WASL. In other words, conditioned upon their characteristics, pupils of a "good" teacher, that is a teacher one standard deviation above average teacher quality, are expected to score between .12 (listening) and .27 (mathematics) WASL standard deviations above pupils of an "average" teacher. Recall, in his work on teacher quality, Rockoff found similar results; the students of a teacher one-standard deviation above average scored between one- and two-tenths of a standard deviation above average on reading and mathematics exams.

One interesting question involves quality correlations by subject. Do teachers who perform well in one subject perform well in another? Table 2 presents correlations between quality measures for all four subject areas. Positive relationships exist between subject areas indicating an affirmative answer; teachers talented in one subject do well in others. The strongest relation occurs between math and reading; instructors teaching math well tend to be good reading teachers. Although all of the correlations are statistically significant, they are not so highly correlated as to suggest that a composite quality measure should be constructed. Thus, the next section models teacher attrition based on each of the four subject-quality measures.

2.3. Teacher attrition estimation

Of the 2293 teachers in this sample, 114 left the profession after the 2001–2002 school year and did not return to a Washington public school in any capacity over either of the subsequent 2 years. The quality distributions of both those that left and remained in the profession are displayed in Fig. 1. For all four quality estimates, there is an apparent tendency for lower quality teachers to exit the profession. As a matter of fact, in only one case (math) did a teacher whose quality is greater than one-half of a WASL standard deviation exit the profession while in every case at least one teacher exited with a quality estimate less than negative one-half of a standard deviation.

Dolton and van der Klaauw (1995, 1999) document large differences in teacher attrition among individuals who exit into non-employment versus into a non-teaching job. Specifically, Dolton and van der Klaauw find that decisions to exit teaching for voluntary reasons (family decisions, taking another job, exiting the labor force, etc.) are sensitive to the explanatory variables in their model (opportunity wages, characteristics of the teacher's school, etc.). Not surprisingly, the same set of

[7] In fact, all regressions reported in Table 1 contain 5 binary variables indicating different levels of reading for fun, 6 binary variables indicating television-watching habits, 3 binary variables indicating the amount of English spoken at home, and 5 racial binary variables. Complete regression results are available from the author upon request.

Table 1
Selected regression estimates of student mathematics 4th grade WASL score

Variable	Description	Reading	Writing	Listening	Math
Student variables					
ITBS subject	Third grade ITBS subject score	0.637***	0.437***	0.341***	0.713***
		(−0.005)	(−0.006)	(−0.006)	(−0.005)
New school	Student is new to school	−0.055***	−0.075***	−0.079***	−0.058***
		(−0.014)	(−0.017)	(−0.019)	(−0.014)
Home computer	Student has computer at home	0.057***	0.047***	0.081***	0.089***
		(−0.013)	(−0.015)	(−0.017)	(−0.011)
Computer use	Student uses a computer for school work	−0.0004	.023**	−0.025**	0.0045
		(−0.009)	(−0.01)	(−0.01)	(−0.008)
Read often	Student often reads for fun	0.130***	0.186***	0.050***	0.084***
		(−0.012)	(−0.014)	(−0.016)	(−0.011)
Hold back	Student was held back a grade	−0.121***	−0.152***	−0.065***	−0.094***
		(−0.014)	(−0.017)	(−0.019)	(−0.012)
English never	English is never spoken in student's home	−0.063***	−0.112***	−0.031***	−0.007
		(−0.016)	(−0.018)	(−0.02)	(−0.014)
Black	Student is Black	−0.095***	−0.049**	−0.114***	−0.159***
		(−0.021)	(−0.024)	(−0.026)	(−0.021)
Asian	Student is Asian	0.156***	0.309***	−0.015	0.012
		(−0.017)	(−0.017)	(−0.02)	(−0.015)
Hispanic	Student is Hispanic	−0.017	−0.019**	−0.183***	−0.074***
		(−0.01)	(−0.01)	(−0.022)	(−0.015)
Male	Student is male	−0.138***	−0.370***	0.115***	−0.081***
		(−0.007)	(−0.008)	(−0.009)	(−0.007)
Building variables					
Building pass	% Of students passing previous year's related WASL subject	0.004***	0.006***	0.002****	0.006***
		(−0.0009)	(−0.0007)	(−0.0005)	(−0.0008)
Peer effects	Average ITBS subject score for student's class	0.023*	0.136***	0.040**	0.002
		(−0.015)	(−0.02)	(−0.017)	(−0.021)
Free lunch	% Of students in building with free/reduced lunch	0.001	−0.0004	0.002**	0.002**
		(−0.0008)	(−0.0008)	(−0.0006)	(−0.0009)
Class size	Size of student's class	−0.0007	−0.0008	0.0009	0.00005
		(−0.002)	(−0.001)	(−0.001)	(−0.001)
Teacher variables					
T. male	Teacher is male	−0.053***	−0.068***	0.001	−0.044***
		(−0.015)	(−0.017)	(−0.013)	(−0.017)
Masters	Teacher holds a master's degree	0.01	0.022	0.014	−0.004
		(−0.013)	(−0.014)	(−0.011)	(−0.015)
Experience	Years of teaching experience	0.006**	0.003	0.007***	0.009***
		(−0.002)	(−0.002)	(−0.002)	(−0.003)
Experience^2	Years of teaching experience squared	−0.0001	−0.00009	−0.0001**	−0.0002***
		(−0.00007)	(−0.00007)	(−0.00007)	(−0.00008)
First year	Teacher in 1st year of teaching	−0.042**	−0.026	−0.007	−0.043
		(−0.02)	(−0.022)	(−0.017)	(−0.025)
R^2		0.66	0.61	0.51	0.68
Slope test	Wald test: slope coefficients equal 0 (χ^2 test)	110,477***	26,325***	198,106***	127,140***
$N_{\text{Observations}}$	# Of student observations	36,056	36,056	36,056	36,056
District	District dummy variables	Yes	Yes	Yes	Yes
N_{Teachers}	# Of teacher coefficients	2293	2293	2293	2293
Teacher effects	Method of teacher effects determination	Random effects	Random effects	Random effects	Random effects

Notes: Dependent variable is student's WASL subject score, which has been normalized so the sample has mean zero and standard deviation 1. Standard errors corrected for clustering by classroom are in parenthesis. *** {**} indicate significance at the 99% {95%} level.

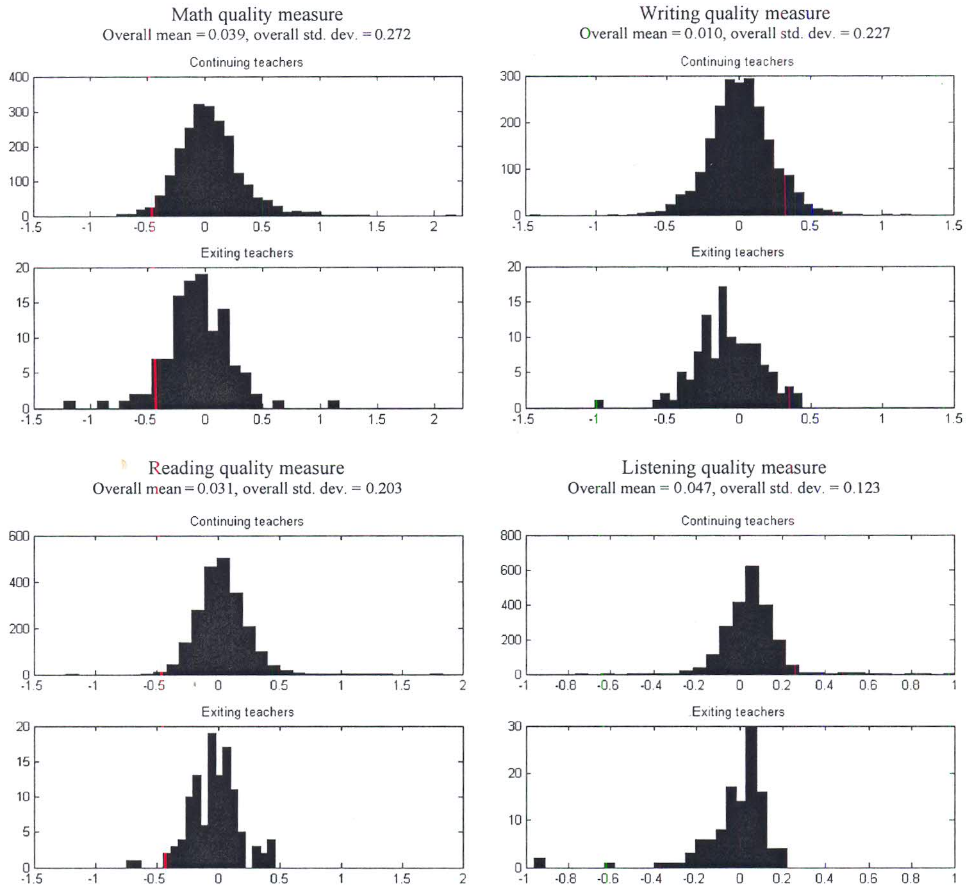

Fig. 1. Teacher quality histograms.

explanatory variables fails to predict involuntary exit. Unfortunately the reason for exit is not observed in this data but, as suggested by Dolton and van der Klaauw, female teachers are both more often secondary wage earners and more likely to be impacted by family transitions, and hence more likely to voluntarily exit. Thus, this paper controls for gender by splitting the sample into male and female teachers. Table 3 provides descriptive statistics of quality and other teacher variables by gender.

The left half of Table 3 focuses on female teachers. Of the 1833 female teachers, 90 (4.9%) exited after the 2001–2002 school year. The average mathematical quality score for women remaining in the profession is .054, a statistically greater estimate than women who exited (−0.057). The same pattern is found in the three other quality measures. The fact that average female teacher quality is greater for those remaining than for those who exit suggests better female instructors remain in the profession. Table 3 also demonstrates that female teachers who exit average lower salaries, fewer years of experience, were on probationary contracts[8] and come from buildings with students who performed worse on the previous year's WASL exam. Further, women leaving teaching are likely to be older and work in larger schools.

The right half of Table 3 provides descriptive statistics for male teachers. Of the 460 male teachers sampled, 24 (5.2%) exited at the end of the school year. Similar to women, men who exit are older and lower quality as measured by the listening scores. Unlike female teachers, no other significant differences in other quality measures or any other variables exist between men who stay or exit the profession.

[8]Washington teachers are classified as being either probationary teachers or teachers under "continuing" contracts. A probationary teacher is a teacher with fewer than 2 years of experience or an experienced teacher with less than 1 year of experience at a new district. Teachers under continuing contracts may not be renewed only for gross negligence or for district financial need.

Table 2
Correlations of teacher quality measures

	$\hat{Q}_{\text{Math}}$	$\hat{Q}_{\text{Reading}}$	$\hat{Q}_{\text{Writing}}$	$\hat{Q}_{\text{Listening}}$
$\hat{Q}_{\text{Math}}$	1.000			
$\hat{Q}_{\text{Reading}}$	0.589***	1.000		
$\hat{Q}_{\text{Writing}}$	0.449***	0.571***	1.000	
$\hat{Q}_{\text{Listening}}$	0.329***	0.415***	0.269***	1.000

***Indicates statistical significance at the 99% level.

In order to control for the many influences on teacher exit, logits with dependent variable equal to one if the teacher left the profession at the end of the school year are estimated. For each gender, four logits are estimated; one each for the four different measures of teacher quality. In order to simplify discussion, each quality measure developed in the first stage regression was normalized with mean zero and variance equal to one. Besides the quality measures, a number of additional independent variables are included. Murnane and Olsen (1989), as well as Dolton and van der Klaauw (1995, 1999) find that teacher attrition increases with potential non-teaching wages. To control for this both the teacher's salary, dummy variables representing the 39 Washington counties in which the teacher's live, as well as dummy variables for the size of community the teacher's building is located are included in the logit models. Yet it is clear that influences other than compensation impact attrition. As a matter of fact, Scafidi, Sjoquist, and Stinebrickner (2002) suggest that teacher earnings only partly explain attrition.[9] Thus, teacher's experience, probationary contract status, race, age, education level, and the size of the teacher's class are included in the logit models. Also, incorporated are building variables: enrollment, percent of students on free or reduced lunch programs, as well as the percent of students who are black, Hispanic, or American Indians. Hanushek, Kain, and Rivkin (2001) suggest that teachers of more talented students tend to remain in education. In order to control for this, two measures of student ability are included: the average score on each ITBS discipline test of the teacher's students and the overall building WASL pass rate from the previous year.

Tables 4 and 5 present logit estimates of the marginal impacts of quality on female and male teacher exits, respectively. Are higher-quality teachers more likely to exit? For females, the answer to this question appears to be no. As a matter of fact, for all four measures of teacher quality, the impact of quality on the probability of female exit is negative. For example, consider the math quality logit regression. A one standard deviation increase in teacher quality (equivalent to .272 standard deviations in student math performance) is expected to diminish attrition by 1.08%. Given the fact that 4.9% of women teachers exited at the end of the 2001–2002 school year, a 1.08% difference represents a 22% decrease in propensity to exit. Two other measures of quality, writing and listening, demonstrate even larger impacts. Women one standard deviation above average writing and listening are 1.11 and 1.23% less likely to exit respectively. All three estimates are measured precisely enough to exclude zero from the 99% confidence interval. Interestingly, this is not the case for the reading quality measure. Although the point estimate of reading quality on attrition is negative, its corresponding confidence interval is large. Recall, in all four cases, these estimates are likely biased towards zero to the extent that the quality variable is measured with error.

Compared to the female results, the impact of quality on male exit is smaller, measured less precisely, and in all four regressions, not significantly different than zero at the 95% level. Dolton and van der Klaauw (1999) suggest a possibility for these findings: involuntary exits typically occur for reasons orthogonal to teacher quality (expiring contracts, retirement, demographic shifts which lead to downsizing, etc.). If men, upon becoming teachers, more frequently exit for involuntary rather than voluntary reasons, then one would not expect quality to predict attrition. This also explains why few other explanatory variables are statistically significant in the male regressions.

Similarities between male and female attrition are apparent when comparing Tables 4 and 5. Not surprisingly, both genders respond to higher salaries by remaining in education, a finding in agreement with those of Dolton and van der Klaauw (1999) and Murnane and Olsen (1989). Further, both genders demonstrate a U-shaped probability of exit with respect to experience. Teachers with both few and many years of experience are more likely to exit than those in the middle of their careers, a finding consistent with Hanushek, Kain, and Rivkin's (2001) work.

Finally, unlike men, women are found to exit if they serve in buildings with less-talented students, as measured by previous WASL results. Surprisingly a woman's exit rate is not correlated with the average ability of her own students as measured by their class ITBS average. Women are also found to be more likely to exit if they serve in larger buildings and if they live in an urban environment (rural area is the omitted variable). Surprisingly, for neither gender, the type of degree earned, race of the teacher, or the size of their most recent class predicts attrition.

[9]In Scafidi, Sjoquist, and Stinebrickner's (2002) work, the authors report that for a sample of female teachers in Georgia, only 5.4% of high school teachers and 3.8% of elementary teachers leaving education earn more on their new job than a first year teacher.

Table 3
Descriptive statistics

Variable	Description	Female teachers			Male teachers		
		Exit		Stay	Exit		Stay
$\widehat{\text{Quality}}_{\text{Math}}$	Quality conditional on student, building, and district characteristics	−0.057	<	0.054	−0.05	=	0.009
		(−0.274)		(−0.275)	(−0.288)		(−0.246)
$\widehat{\text{Quality}}_{\text{Reading}}$	Quality conditional on student, building, and district characteristics	−0.022	<	0.044	−0.065	=	−0.003
		(−0.208)		(−0.207)	(−0.197)		(−0.18)
$\widehat{\text{Quality}}_{\text{Writing}}$	Quality conditional on student, building, and district characteristics	−0.079	<	0.028	−0.085	=	−0.036
		(−0.233)		(−0.227)	(−0.215)		(−0.211)
$\widehat{\text{Quality}}_{\text{Listening}}$	Quality conditional on student, building, and district characteristics	−0.022	<	0.052	−0.017	<	0.045
		(−0.183)		(−0.122)	(−0.137)		(−0.101)
Salary	Annual salary, $1000s	42.324	<	47.526	51.434	=	49.156
		(−14.502)		(−10.61)	(−16.96)		(−10.94)
Experience	Years of experience	10.926	<	12.773	16.12	=	14.173
		(−12.083)		(−9.14)	(−13.98)		(−10.26)
Probationary	Equal to 1 if the teacher is on a probationary contract	0.277	>	0.124	0.25	=	0.117
		(−0.45)		(−0.33)	(−0.44)		(−0.321)
Age > 55	Equal to 1 if age is greater than 55	0.211	>	0.108	0.291	>	0.112
		(−0.41)		(−0.311)	(−0.464)		(−0.316)
Masters of education	Equal to 1 if highest degree is masters of education	0.444	=	0.546	0.5	=	0.578
		(−0.499)		(−0.498)	(−0.51)		(−0.494)
Other masters	Equal to 1 if highest degree is non-education masters	0.066	=	0.049	0.083	=	0.064
		(−0.25)		(−0.217)	(−0.282)		(−0.245)
Black	Equal to 1 if teacher is black	0.011	=	0.014	0	=	0.009
		(−0.105)		(−0.116)			(−0.095)
Hispanic	Equal to 1 if teacher is Hispanic	0.011	=	0.018	0	=	0.028
		(−0.105)		(−0.132)			(−0.163)
Building pass	% Of building's students who passed previous WASL	26.281	<	29.681	22.758	=	27.124
		(−12.77)		(−13.77)	(−12.91)		(−13.32)
Pupils per teacher	Teacher's class size	18.423	=	18.533	18.021	=	18.235
		(−2.47)		(−2.33)	(−2.54)		(−2.24)
Building enrollment	Total student enrollment in teacher's building	505.92	>	477.89	481	=	453.73
		(−164.89)		(−144.5)	(−146.51)		(−141.32)
Free/reduced lunch	% Of building's students using free/reduced lunch	39.046	=	36.76	43.788	=	39.465
		(−23.89)		(−22.39)	(−24.85)		(−22.2)
Unemployment rate	Unemployment rate in teacher's county	6.676	=	6.618	6.816	=	6.795
		(−2.13)		(−1.95)	(−1.93)		(−2.063)
County wages	Average wages per insured worker in county, $1000 s	35.878	=	34.76	34.53	=	33.932
		(−8.92)		(−8.17)	(−7.93)		(−8.29)
Large city	= 1 if teacher lives in area with population > 250,000	0.033	=	0.043	0.041	=	0.066
		(−0.18)		(−0.204)	(−0.204)		(−0.249)
Midsize city	= 1 if teacher lives in area with 25,000 < population < 250,000	0.211	>	0.149	0.166	=	0.188
		(−0.41)		(−0.356)	(−0.38)		(−0.391)
N	Number of teacher observations	90		1743	24		436

Notes: > < indicate statistical difference at the 95% level. Standard errors in parenthesis.

Table 4
Marginal logit estimates of female teacher attrition

Variable	Logit 1	Logit 2	Logit 3	Logit 4
$\hat{\text{Quality}}_{\text{Math}}$	−0.0108***			
	(−0.003)			
$\hat{\text{Quality}}_{\text{Reading}}$		−0.0058		
		(−0.004)		
$\hat{\text{Quality}}_{\text{Writing}}$			−0.0111***	
			(−0.003)	
$\hat{\text{Quality}}_{\text{Listening}}$				−0.0123***
				(−0.003)
Salary	−0.0017**	−0.0017**	−0.0015**	−0.0017**
	(−0.0007)	(−0.0007)	(−0.0006)	(−0.0007)
Experience	−0.0036***	−0.0039***	−0.0041***	−0.0032**
	(−0.002)	(−0.002)	(−0.001)	(−0.016)
Experience^2	0.00013***	0.00014***	0.00014***	0.00011***
	(−0.00004)	(−0.00004)	(−0.00005)	(−0.00004)
Probationary	−0.001	−0.0005	−0.002	−0.002
	(−0.01)	(−0.011)	(−0.01)	(−0.009)
Age > 55	0.051**	0.052**	0.051**	0.049**
	(−0.026)	(−0.027)	(−0.026)	(−0.025)
Masters of education	0.006	0.006	0.005	0.007
	(−0.008)	(−0.008)	(−0.008)	(−0.007)
Other masters	0.01	0.012	0.008	0.014
	(−0.02)	(−0.022)	(−0.019)	(−0.021)
Building pass	−0.0008**	−0.0008***	−0.0007**	−0.0008**
	(−0.0003)	(−0.0003)	(−0.0003)	(−0.0003)
Pupils per teacher	−0.0018	−0.002	−0.0018	−0.0022
	(−0.001)	(−0.002)	(−0.002)	(−0.002)
Incoming ITBS	0.013	0.01	0.008	0.01
	(−0.007)	(−0.007)	(−0.007)	(−0.007)
Building enrollment	0.00003	0.00004	0.00004	0.00004
	(−0.0002)	(−0.00003)	(−0.00003)	(−0.00002)
Building free/reduced lunch	−0.0002	−0.0002	−0.0003	−0.0002
	(−0.0002)	(−0.0002)	(−0.0002)	(−0.0002)
Large city	0.0228**	0.0257**	0.0238**	0.0245**
	(−0.01)	(−0.011)	(−0.01)	(−0.01)
Midsize city	0.0430**	0.0492**	0.0455**	0.0450**
	(−0.02)	(−0.021)	(−0.02)	(−0.02)
N	1833	1833	1833	1833
Chi square test of 0 slopes	98.84***	90.53***	99.82***	104.31***
County dummy variables present?	Yes	Yes	Yes	Yes
% Of attritions correctly predicted	64.8	69.2	67.4	62.6
% Of continuers correctly predicted	78.9	75.8	76.9	74.8

Notes: Standard errors of quality variables adjusted for measurement error using Gawande's (1997) method. Robust standard errors in (). *** {**} indicate significance at the 99% {95%} [90%] level. The coefficients on the percent of a building's students who are black, Hispanic, or American Indian, as well as the constant, are not reported and are available from the author by request. Incoming ITBS is the classroom average ITBS discipline score matched to the WASL discipline.

2.4. *Alternative specifications*

One concern regarding these findings is their sensitivity to both different methods of quality measurement and to alternative techniques of interacting quality with attrition. Table 6 presents the impact of quality on attrition from combinations of alternative specifications of both Eqs. (1) and (3). For ease of comparison, the first row of Table 6 summarizes the results of Tables 4 and 5.

Table 5
Marginal logit estimates of male teacher attrition

Variable	Logit 5	Logit 6	Logit 7	Logit 8
$\widehat{\text{Quality}}_{\text{Math}}$	−0.0031			
	(−0.006)			
$\widehat{\text{Quality}}_{\text{Reading}}$		−0.0035		
		(−0.0022)		
$\widehat{\text{Quality}}_{\text{Writing}}$			−0.0038	
			(−0.0025)	
$\widehat{\text{Quality}}_{\text{Listening}}$				−0.0102*
				(−0.0059)
Salary	−0.0013**	−0.0013**	−0.0013**	−0.0011**
	(−0.0006)	(−0.0006)	(−0.0006)	(−0.0006)
Experience	−0.008**	−0.008***	−0.008**	−0.007**
	(−0.003)	(−0.003)	(−0.003)	(−0.003)
Experience^2	0.0002**	0.0002***	0.0002**	0.0002**
	(−0.00009)	(−0.00009)	(−0.00009)	(−0.00009)
Probationary	0.009	0.008	0.009	0.008
	(−0.028)	(−0.028)	(−0.027)	(−0.026)
Age > 55	0.024	0.021	0.024	0.029
	(−0.033)	(−0.033)	(−0.033)	(−0.036)
Masters of education	−0.014	−0.013	−0.014	−0.011
	(−0.015)	(−0.015)	(−0.015)	(−0.014)
Other masters	−0.014	−0.014	−0.013	−0.009
	(−0.016)	(−0.016)	(−0.015)	(−0.018)
Building pass	−0.0006	−0.0006	−0.0006	−0.0003
	(−0.0006)	(−0.0005)	(−0.0006)	(−0.0006)
Incoming ITBS	−0.003	−0.008	−0.007	−0.009
	(−0.015)	(−0.014)	(−0.015)	(−0.014)
Pupils per teacher	−0.0012	−0.0013	−0.001	−0.0011
	(−0.003)	(−0.003)	(−0.003)	(−0.003)
Building enrollment	0.00005	0.00004	0.00005	0.00004
	(−0.00005)	(−0.00005)	(−0.00005)	(−0.00005)
Building free/reduced lunch	−0.0002	−0.0002	−0.0002	−0.0001
	(−0.0004)	(−0.0004)	(−0.0003)	(−0.0004)
Large city	0.021	0.02	0.021	0.02
	(−0.02)	(−0.02)	(−0.02)	(−0.02)
Midsize city	−0.018	−0.017	−0.019	−0.016
	(−0.016)	(−0.016)	(−0.016)	(−0.016)
N	460	460	460	460
Chi square test of 0 slopes	34.22***	34.60***	34.57***	37.81***
County dummy variables present?	Yes	Yes	Yes	Yes
% Of attritions correctly predicted	75	75	75	79.1
% Of continuers correctly predicted	77.5	78.2	77.2	78.9

Notes: Standard errors of quality variables adjusted for measurement error using Gawande's (1997) method. Robust standard errors in (). *** {**} indicate significance at the 99% {95%} [90%] level. The coefficients on the percent of a building's students who are black, Hispanic, or American Indian, as well as the constant, are not reported and are available from the author by request. Incoming ITBS is the classroom average ITBS discipline score matched to the WASL discipline.

Because wide variation of building characteristics exist within some districts, one concern with the measure of quality used in Eq. (1) is that of omitted building variables. For instance, the largest school district represented in this sample, Seattle, contains 115 buildings that offer the fourth grade some of which have no students on free lunch and at least one building with grater than 85% participation in the free lunch program. Further, teaching philosophy and student achievement is likely influenced by building principals and building-specific variables such as parental involvement and social-demographic influences. By including fixed effects

Table 6
Alternative specifications of teacher quality and attrition relationship

	Quality estimation specification (Equation 1)	Teacher attrition specification (Equation 3)	Math quality marginal impact on attrition	Reading quality marginal impact on attrition	Writing quality marginal impact on attrition	Listening quality marginal impact on attrition
(1)	Teacher random effects with district dummy variables	Logit with county dummy variables	Women = −0.0108*** Men = −0.0031	Women = −0.0058 Men = −0.0035	Women = −0.0111*** Men = −0.0038	Women = −0.0123*** Men = −0.0102*
(2)	Teacher random effects with district and building dummy variables	Logit with county dummy variables	Women = −0.0064** Men = −0.0065	Women = −0.0007 Men = −0.0048	Women = −0.0056** Men = −0.0029	Women = −0.0086** Men = −0.0116*
(3)	Teacher random effects with district and building dummy variables	Logit with county and district dummy variables	Women = −0.0072** Men = −0.0031	Women = −0.0004 Men = −0.0023	Women = −0.0062** Men = −0.0014	Women = −0.0091** Men = −0.0078*
(4)	Teacher random effects with district and building dummy variables	Building random effects logit with county and district dummy variables	Women = −0.0058** Men = −0.0019	Women = −0.0011 Men = −0.0024	Women = −0.0063*** Men = −0.0027	Women = −0.0034* Men = −0.0065*
(5)	Teacher fixed effects with district dummy variables	Logit with county dummy variables	Women = −0.0131*** Men = −0.0028	Women = −0.0006 Men = −0.0096	Women = −0.0069** Men = −0.0005	Women = −0.0073** Men = −0.0154
(6)	Teacher fixed effects with district dummy variables	Logit with county and district dummy variables	Women = −0.0086** Men = −0.0002	Women = −0.0015 Men = −0.0028	Women = −0.0082*** Men = −0.0028	Women = −0.0081*** Men = −0.0082
(7)	Teacher fixed effects with district dummy variables	Building random effects Logit with county and district dummy variables	Women = −0.0087** Men = −0.0027	Women = −0.0018 Men = −0.0090	Women = −0.0066** Men = −0.0014	Women = −0.0075** Men = −0.0153*
(8)	Teacher fixed effects with district and building dummy variables	Building random effects Logit with county and district dummy variables	Women = −0.0053** Men = −0.0018	Women = −0.0037 Men = −0.0058	Women = −0.0041** Men = .0013*	Women = −0.0023* Men = −0.0007

All coefficients represent marginal impact of quality on teacher attrition. *** {**} indicate significance at the 99% {95%} [90%] level using Gawande's (1997) method of adjusting the standard errors of generated variables. When building dummy variables are used in the quality estimation models the building pass and free lunch variables were omitted. Teacher fixed effects model exclude teacher gender, experience, education, class size, and peer effects variables.

for buildings in Eq. (1), the generated quality measure is conditioned upon building characteristics invariant across students such as the quality of principals. The second row of Table 6 reports the impacts of quality on attrition estimated from Eq. (3) using quality measures generated by Eq. (1) augmented by building fixed effects.[10] Although the point estimates decrease in magnitude by about 40%, similar results to the original model are found. Using this alternative measure of quality; women with higher quality scores (other than reading) are likely to remain as teachers while the impact of quality on male attrition is insignificant.

A related concern is that different districts have policies that discourage teacher exit. For instance, some districts lighten a teacher's workload by hiring considerable numbers of paraprofessional teachers while other districts hire very few supplementary instructors. This difference in work conditions may impact teacher attrition. If these policies are related to quality, then the logit model explaining teacher attrition will be biased. In order to control for this possibility, district dummy variables were added to the logit models of Eq. (3). The resulting estimates of quality on attrition are listed in the third row of Table 6 and again, no qualitative differences are found when compared to the original model.

Another possibility is that teacher attrition depends upon unobservable building characteristics. At the building level, strong principals may generate greater annual gains among students than buildings with less able principals. If strong administrators also tend to retain teachers, then the methods used previously would overstate teacher quality and incorrectly find better teachers are less likely to leave. Further, the variety of fourth grade situations in Washington buildings varies greatly. For instance, one school building in the Seattle school district operates 11 different 4th grade classrooms while another building in rural Washington operates a 4th grade classroom with 3 students. In order to control for unobserved variation by school building that influences on attrition, the fourth row of Table 6 employs a building-level random-effects logit specification. Combined with the quality measure generated with building dummy variables, the results reported in the fourth row of Table 6 most comprehensively control for unobserved building- and district-level heterogeneity. Although the magnitudes estimated in this procedure are smaller than the previous three, the writing, listening, and mathematics estimates remain both statistically and economically important for women teachers. A one-standard deviation increase in math and writing quality decreases attrition probabilities by .58% and .63%, respectively.

In order to check for robustness across panel estimation techniques, the final four rows of Table 6 report the impact of quality on attrition when equation one is estimated with fixed rather than random effects.[11] This technique was employed by Rockoff (2004) to measure teacher quality. In all fixed effects specifications, the same pattern as the random effects specifications was found: high ability women are less likely to exit while no significant pattern for men exists.

One consistent result across specifications involving women is the insignificance of reading quality on attrition. In light of the significance of the other quality variables, this finding is surprising. One potential explanation is that reading quality suffers from measurement error. This is suggested by Table 3's histogram of reading quality that demonstrates a small number (9) of high ability teachers who chose to exit. To check for this possibility, the high-ability instructors who exited were excluded from the sample and the building random effects specification re-estimated. The (unreported) results are qualitatively the same; reading quality does not impact attrition for either men or women.[12]

3. Discussion

Teacher quality and student performance are inherently important issues to public education. Public perception is that schools lose their best teachers to outside opportunities while retaining the least competent. The evidence presented in this paper suggests otherwise. Specifically, fourth grade female teachers are less likely to leave teaching if their students perform above expectations after controlling for student, building, and teacher characteristics. Depending upon the specification used, a female teacher one standard deviation above her peers in mathematics quality is between .58% and 1.3% less likely to exit at the end of the school year. Similar, if not slightly larger, estimates are found for both writing and listening quality measures. Over the time of this study, 4.9% of female teachers left the profession, thus a .58% change in the

[10]Because each building is represented by a dummy variable in this construction, it was necessary to exclude the two measures in Eq. (1), which did not vary within a building: the percent of students on free/reduced lunch and the percent of previous students who passed the WASL.

[11]Estimating teacher quality through fixed effects precludes the use of any measurable teacher variable that does not vary within a classroom. Thus teacher's experience, gender, education, class size, and peer effects are excluded from Eqs. (1) and (2) in these specifications.

[12]In fact a second specification involved excluding all teachers, whether they exited or remained in the profession, who scored more or less than three standard deviations from the quality mean. Again, no qualitative differences from the results reported in Table 6 were found.

probability of exit represents nearly a 12% change in attrition potential. To put this in context, Table 4 reports an additional $1000 salary is estimated to diminish female attrition by .17%—thus a .58% change in the probability of exit is equivalent to raising annual salaries by $3400—equivalent to a 7% annual raise for the average teacher. Given the likelihood of both measurement error and an upward bias in the quality estimates, there is reason to believe that these estimates represent a lower bound on the impact of quality on attrition.

The finding that high-ability women are less likely to exit is interesting in light of recent research by Hoxby and Leigh (2004) and Corcoran, Evans, and Schwab (2004), who argue that the average quality of new female teachers has declined over time as women have succeeded in employment fields traditionally reserved for men. Although my work cannot address quality changes over long time periods nor does it focus on entrants into teaching, the fact that high quality female teachers are less likely to exit the profession mitigates this decline.

A number of reasons for explaining a negative relationship between quality and attrition are possible. First, it is likely that teachers who perform well receive valuable intangible benefits. Teachers value the esteem of their colleagues and may increase their chances of choosing future students, buildings, and districts. It is possible that better performing teachers receive promotions, class release time, or professional development opportunities. It is also likely that individuals with the inclination to teach will be more able instructors. If that inclination is such that these individuals would require a pay premium to leave teaching and if non-education employers refuse to pay this premium, then higher-quality teachers would be less likely to exit. Of course individuals attracted to education likely value high performance in their students and would be less likely to quit.

Other explanations for this paper's findings are possible. For instance, if truly high quality teachers refuse to "teach to the test," and low quality teachers do, then it is possible that low quality teachers are incorrectly identified as high quality using a test-based measure. If the truly low quality teachers have fewer job opportunities, are less likely to exit teaching, and more likely to "teach to the test", then one would expect to find a negative relationship between attrition and a test-based measure of quality. However, presumably men have the same temptation to teach to the test as women yet no significant negative relationship was found for men suggesting that the results were not driven by teaching to the test.

The findings in this paper may also be influenced by the grade level examined. The extent to which the skills that make a successful fourth grade teacher marketable in other jobs is unknown. If these skills are specific to the teaching profession and successful teachers spend effort developing these skills, then it is not surprising that they are less likely to exit. Having invested in teaching-specific skills, perhaps the high quality fourth grade teacher has invested in fewer general skills valued by other employers. This is likely to be less of a factor at higher-grade levels where teachers specialize in an academic field of study, which is potentially more valued by employers. Thus, one must be careful extending these results to higher grades.

The grade level examined possibly explains why women of higher reading quality demonstrate insignificant attrition proclivities whereas other quality measures are negatively correlated with attrition. If reading instruction is the prime goal of the fourth grade, then individuals teaching reading well may be more attractive to private schools or social agencies dealing with children of that age. These added opportunities to exit may be counterbalanced by the intangible benefits of being a good reading instructor with the end result being an insignificant impact of reading quality on attrition. If math, writing, and listening are skills less emphasized at the fourth grade, then teachers succeeding in these subjects would not receive as many opportunities to leave public schools.

This paper documents a large disparity in teacher attrition when sorted by gender. Generally, higher-quality female teachers are less likely to leave whereas quality is not a predictor of male exits. As a matter of fact, other than salary and experience, little else predicts male teacher exit. Dolton and van der Klaauw (1999) suggest a possible reason: rather than choosing to stay at home with a family or exiting to follow a spouse who changes jobs, men are more likely to exit teaching only upon retirement or the termination of a contract. On the other hand, women, who are likely to be second-wage earners, have more latitude to leave teaching. With this latitude, women are more sensitive to the intangible benefits of quality and job conditions.

Finally, the econometric methodology employed by this paper has possible extensions in the education literature as well as applied work in other fields. The interaction between salary and quality is of obvious interest as is the decision of higher-quality teachers to leave districts for better schools. As mandated by the NCLBA, failing schools are required to replace their teachers. A measure of teacher quality will help explain if those schools are failing because of poor teachers or because those schools draw from an underachieving group of students.

Acknowledgements

The author would like to thank Paul Storer, Yvonne Durham, John Robst, Darryl Getter, and two anonymous referees for their helpful comments. The data used were generously provided by Pete Bylsma of Washington's Office of the Superintendent of Public Instruction.

References

Amemiya, T. (1978). On a two-step estimation of a multivariate logit model. *Econometrica*, *8*(1), 13–21.

Corcoran, S., Evans, W., & Schwab, R. (2004). Changing labor-market opportunities for women and the quality of teachers, 1957–1992. *American Economic Review*, *94*(2), 236–240.

DeVaro, J., & Lacker, J. (1995). Errors in variables and lending discrimination. Federal Reserve Bank of Richmond. *Economic Quarterly*, *81*(3), 19–32.

Dolton, P., & van der Klaauw, W. (1995). Leaving teaching in the UK: a duration analysis. *Economic Journal*, *105*(429), 431–444.

Dolton, P., & van der Klaauw, W. (1999). The turnover of teachers: a competing risks explanation. *Review of Economics and Statistics*, *81*(3), 543–550.

Gawande, K. (1997). Generated regressors in linear and nonlinear models. *Economics Letters*, *54*(2), 119–126.

Greene, W. (2000) *Econometric Analysis* (4th ed.). Upper Saddle River, NJ.

Hanushek, E., Kain, J., & Rivkin, S. (1998). *Teachers, schools, and academic acheivement*. NBER Working Paper 6691.

Hanushek, E., Kain, J., & Rivkin, S. (2001). *Why public schools lose teachers*. NBER Working Paper 8599.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161.

Hoxby, C., & Leigh, A. (2004). Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States. *American Economic Review*, *94*(2), 236–240.

Murnane, R., & Olsen, R. (1989). The effects of salaries and opportunity costs on duration in teaching: evidence from Michigan. *Review of Economics and Statistics*, *71*(2), 347–352.

Rockoff, J. (2004). The impact of individual teachers on student achievement: evidence from panel data. *American Economic Review*, *94*(2), 247–252.

Scafidi, B., Sjoquist, D., & Stinebrickner, T. (2002). *Where do teachers go*? Working Paper.

Stoddard, C. (forthcoming). Adjusting teacher salaries for the cost of living: the effect of salary comparisons and policy conclusions. *Economics of Education Review*.

Further reading

Hanushek, E. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, *24*(3), 1131–1177.

Meryman, R. (1962). How we drive teachers to quit. *Life*, *53*, 104–106.