Women, the Labor Market, and the Declining Relative Quality of Teachers

Sean P. Corcoran
William N. Evans
Robert M. Schwab

Abstract

School officials and policymakers have grown increasingly concerned about their ability to attract and retain talented teachers. A number of authors have shown that in recent years the brightest students—at least those with the highest verbal and math scores on standardized tests—are less likely to enter teaching. In addition, it is frequently claimed that the ability of schools to attract these top students has been steadily declining for years. There is, however, surprisingly little evidence measuring the extent to which this popular proposition is true. We have good reason to suspect that the quality of those entering teaching has fallen over time. Teaching has for years remained a predominately female profession; at the same time, the employment opportunities for talented women outside teaching have soared. In this paper, we combine data from five longitudinal surveys of high school graduates spanning the classes of 1957 to 1992 to examine how the propensity for talented women to enter teaching has changed over time. While the quality of the average new female teacher has fallen only slightly over this period, the likelihood that a female from the top of her high school class will eventually enter teaching has fallen dramatically. © 2004 by the Association for Public Policy Analysis and Management.

> "The quality of teachers has been declining for decades, and no one wants to talk about it.... We need to find a more powerful means to attract the most promising candidates to the teaching profession."
>
> —Harold O. Levy, chancellor of the New York City Public Schools in "Why the Best Don't Teach" (*New York Times,* September 9, 2000).

INTRODUCTION

Teacher shortages and concerns over the quality of the teaching force have become perennial issues in the United States. With each passing year, school officials bemoan their inability to attract top candidates into the teaching profession, and the debate over how best to attract and retain talented, better-qualified teachers seems to intensify. A popular explanation for these frustrations, outlined in Temin (2002),

Manuscript received August 2003; review complete November 2003; accepted December 2003

Journal of Policy Analysis and Management, Vol. 23, No. 3, 449–470 (2004)

Published by Wiley Periodicals, Inc. Published online in Wiley InterScience (www.interscience.wiley.com)
DOI: 10.1002/pam.20021

points to the remarkable gender desegregation of the labor market since 1960.[1] Schools that once found a captive labor pool in college-educated women are today forced to compete with a diverse array of professions, with the best and brightest believed to be least likely to enter teaching. This belief has been the impetus behind many recent policy measures—salary increases, subsidized housing, and relaxed testing and course requirements, to name a few—each intended to increase the attractiveness of teaching relative to other occupations.

This frustration over the quality of the teaching force comes amidst a growing body of evidence that shows that certain measures of teacher quality—in particular, their verbal and mathematical skills—are strongly related to student outcomes.[2] The persistence of this finding stands in sharp contrast to the continuing debate over the importance of other measured inputs into the production of education (notably, per-pupil expenditure and class size).[3] Recognizing the apparent importance of attracting quality teachers, the federal government responded in 2001 with an unprecedented $2.9 billion program specifically targeted toward improving teacher quality.[4]

While the hypothesis that gender desegregation of the labor market has affected the quality of teachers appears to be widely accepted, there is surprisingly little evidence measuring the extent to which it is true. This is largely due to a lack of data, as recognized in a recent paper by Podgursky, Monroe, and Watson (in press):

> Economists have hypothesized a secular decline in teacher quality as a consequence of rising non-teaching earnings and job opportunities for high ability women. In this view, public schools benefited from the occupational crowding of women into the teaching profession. Unfortunately, time-series data on teacher quality are not available to directly test this hypothesis.

In this paper, we address this issue by combining data from five longitudinal surveys of high school students spanning more than four decades to examine two interrelated trends: how the propensity for women with high verbal and mathematical skills to become teachers has changed over time; and how, as a consequence, the skill distribution of teachers has evolved since 1964. The inclusion of two datasets that date prior to 1965 affords us the opportunity to provide some evidence as to how this relationship between academic ability and entry into teaching has changed over a long period of labor market desegregation.[5]

[1] Temin argues that our history of widening opportunities for women since 1960 has created conditions for the existence of multiple teacher pay-quality equilibria, and that U.S. schools are currently stuck in a low pay-low teacher quality equilibrium. Prior to 1960, he argues, multiple equilibria were not possible as women at all levels of ability were confined to a small number of occupations.

[2] See, for example, Hanushek (1970, 1971), Ehrenberg and Brewer (1995), Ferguson and Ladd (1996), and Hanushek, Kain, and Rivkin (1998). Wayne and Youngs (2003) and Rice (2003) review recent literature on the relationship between teacher characteristics and student achievement.

[3] Hanushek (1986, 1996) and Krueger (2002) provide contrasting views on this issue.

[4] Title II of the Elementary and Secondary Education Act (ESEA), or “No Child Left Behind.” Local education agencies receiving money under Title I of ESEA must also ensure that all teachers hired and supported using ESEA funds are “highly qualified,” a criteria based largely on educational attainment and subject matter knowledge.

[5] We will be studying changes in the reduced-form relationship between academic ability and entry into teaching without explicitly investigating the channels through which this relationship may have been shaped. In other words, we will be observing the outcome of both demand and supply decisions in the market for teachers. Consequently, our results here should be interpreted as descriptive evidence on the academic ability of those individuals who actually identified themselves as teachers during this period. In a recent paper, Bacolod (2003) relates changes in relative teacher salaries to changes in the quantity and quality of teachers over time. Hoxby and Leigh (in press) investigate two hypotheses for the changing distribution of teachers—alternative labor market opportunities and unionization.

WOMEN, THE LABOR FORCE, AND THE TEACHING PROFESSION, 1964–2000

As is well known, women's relationship with the labor market fundamentally changed after 1960. (See Goldin [2004] for an eloquent discussion of this dramatic change.) Labor-force participation among women aged 25–34 more than doubled between 1964 and 2000; at the same time, the fraction of young women with a 4-year college degree tripled (see Table 1). By comparison, the fraction of young men with a 4-year degree rose only 50 percent over the same period. These trends were likely aided in part by key legislative movements of the early 1960s and 1970s that significantly altered the landscape of occupational opportunity for women.[6]

Many highly skilled professions became considerably less gender segregated over this period (see, e.g., Blau, Simpson, and Anderson, 1998; Beller, 1992). In Table 1 we compute a gender representation index for two traditionally high-skill professions—medicine and law—using data on men and women aged 25–34 from the 1968–2000 March Current Population Surveys (CPS). This index is calculated as the proportion female in each occupation divided by the female share of the labor force; a value less than one suggests that women are underrepresented in the occupation, relative to their representation in the labor force. (This and other methods of measuring occupational segregation and concentration are discussed in Siltanen, Jarman, and Blackburn, 1995.) In the mid-1960s this index was 0.33 in medicine, and was virtually zero in law (0.08). In less than thirty years, the value of this index for both occupations had nearly reached one (0.89 and 0.87, respectively).

Table 1. Labor market characteristics, selected years, 1964–2000.

	1964	1970	1980	1990	2000
Fraction of young females (age 25–34) with:					
At least a high school degree	67.4	73.3	84.3	86.5	87.0
At least a 4-year college degree	8.9	12.0	20.7	22.8	27.8
In the labor force	37.2	45.7	65.4	73.2	80.0
Index of gender representation (age 25–34)					
Physicians	0.33	0.20	0.47	0.69	0.89
Lawyers	0.08	0.20	0.44	0.60	0.87
Teachers	2.33	1.83	1.67	1.64	1.62
Public elementary teachers (millions)	1.0	1.1	1.2	1.4	1.7
Public secondary teachers (millions)	0.6	0.9	1.0	1.0	1.1
Public school enrollment (millions)	42.2	45.9	40.9	41.2	47.2
Fraction of young teachers (age 25–34) that are female:					
All	70.8	70.0	72.4	74.2	74.5
Primary	—	85.1	84.6	86.0	84.1
Secondary	—	46.6	54.4	53.0	54.2

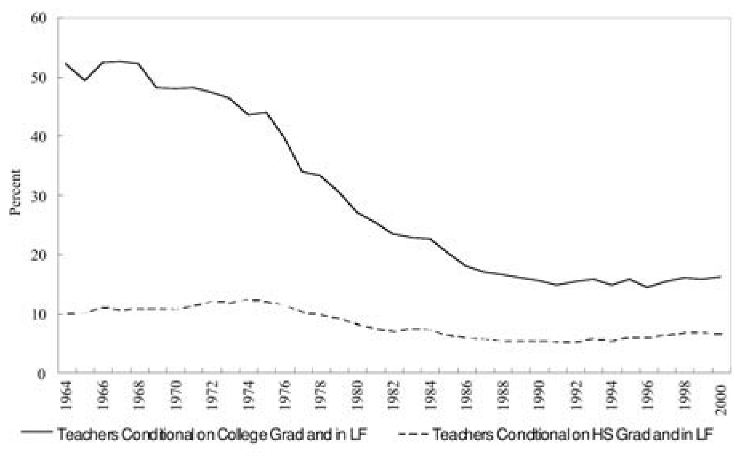
Source: March Current Population Survey 1964–2000, except teacher counts and enrollments, taken from the Digest of Education Statistics, 2000 and 2002. First teacher count figure for 1960, first enrollment is from 1965. Gender representation index is calculated from data originating in 1968.

[6] The Equal Pay Act of 1963 and Title VII of the Civil Rights Act of 1964 were among the most notable, replacing "protective labor laws" and explicitly outlawing employer discrimination on the basis of sex. See Lloyd and Niemi (1979) for a discussion.

Despite striking shifts in the gender composition of professions such as medicine and law, the gender composition of new teachers has remained roughly constant. Women continue to dominate new recruits into the teaching profession, comprising roughly three quarters of teachers aged 25–34—a fraction that has remained virtually unchanged since 1964 (Table 1). The proportion is higher for elementary teachers (84 to 86 percent) than secondary teachers (46 to 54 percent), although the fraction has risen somewhat over time for secondary teachers. A relatively constant fraction of teachers who are women, together with steadily rising female labor force participation, have combined to cut the gender representation index for teachers by more than half, from 2.3 to 1.6 (Table 1).

While teaching remains a predominantly female occupation, the profession itself has significantly diminished in importance as a career path for female college graduates. Nowhere is this as evident as in the plummeting percentage of (working) female graduates who identify themselves as teachers (Figure 1). In 1964, more than half of working female college graduates were teachers—by 2000, this percentage had dropped to 15 percent. While this decline in the fraction of graduates choosing to teach can be attributed largely to changes in the denominator (the enormous rise in college completion among women), one thing in this picture is clear: Conditional on working, of those women who acquired a college education in the 1960s, most went into teaching; of those completing college today, most do not. This raises the question—how has the skill distribution of those women who do choose to teach changed over time?

Occupational desegregation and a significantly increased representation of females in highly skilled professions, such as medicine and law, do not necessarily imply a reduction in the quality of teachers. Growth in the supply of female college graduates far outpaced the growth in teachers over the 1960–2000 period (see Table 1). While the stock of teachers increased by 75 percent between 1960 and 2000, the fraction of young women holding college degrees tripled. Further, this rise in college completion appears to have occurred across the ability distribution, not exclusively at the bottom. By our calculation, women who scored near the top of their



Source: March CPS 1964–2000

Figure 1. Percentage of working females (age 25–34) who are teachers.

high school class were nearly twice as likely to earn a college degree in 2000 vs. 1964.[7] With a simultaneous rise in female selection into high human capital occupations *and* a rise in the supply of high-ability women available to all professions, it remains an empirical question as to how these competing trends have affected the quality of teachers.

MEASURING TEACHER QUALITY

Ideally, teacher quality would be measured using a multi-dimensional vector of those characteristics that are positively associated with outputs of the educational process. This vector likely includes many attributes—patience, creativity, or communication skills, for example—that are difficult to measure and are practically unobserved by the statistician. Indeed, recent research by Hanushek, Kain, and Rivkin (1998) and Rockoff (in press) suggests that unobserved characteristics of teachers (as captured by teacher fixed effects) are perhaps more important for student performance than most observed characteristics, such as experience, certification, and degree attainment.[8]

In this paper, we use a measure of relative cognitive ability—cohort-specific rankings based on scores on a test of verbal and mathematical skills—as our measure of quality. While we acknowledge the obvious limitations of this single-dimensional measure of quality, it seems to us difficult to believe that measures of a teacher's verbal and mathematical ability are *not* an important dimension of teacher quality. After all, these are tests of skills that teachers are expected to cultivate in their own students.

In fact, there is a growing literature that finds that measures of cognitive ability—particularly teachers' verbal ability scores—are among the most important measurable determinants of student achievement (Rice, 2003). This literature suggests that students of teachers who scored higher in the distribution of standardized test takers tend to perform better than students with teachers who ranked lower in the distribution. For example, Ehrenberg and Brewer (1995) find that a one-half standard deviation increase in the verbal aptitude score of white female teachers would have raised the synthetic gain scores of white elementary students in the 1966 Coleman Report data by 4 to 8.5 percent. Likewise, Ferguson and Ladd (1996) find that a one standard deviation increase in teachers' ACT composite scores (in the state of Alabama) would have resulted in a one-tenth standard deviation increase in student reading scores from 3rd to 4th grade (comparable to about one-half the black-white test score gap in urban areas during this time period).

What a teacher's relative test score measures is less clear, but this literature has illustrated that these scores do indeed capture something (whether specific skills or general intelligence) that is important in explaining the academic achievement of their students. It is important to bear in mind, however, that standardized test scores are a measure of relative quality and not absolute quality. While the evidence suggests that students have better outcomes with teachers of greater relative quality, it is entirely possible that the absolute level of human capital in the teaching profession has risen over time. Unfortunately, good measures of absolute quality are not available.

[7] Results available from the authors upon request.
[8] See Rice (2003) for a recent survey. It is important to note that in neither of the studies were measures of academic ability observed.

EXISTING EVIDENCE ON THE QUALITY OF TEACHERS

A number of cross-sectional studies have shown that the verbal and math abilities of college graduates entering teaching in the 1970s and 1980s did not compare favorably to that of their peers who chose alternative professions. Many of these studies rely on comparisons between mean SAT or ACT scores of examinees intending to major in education with those who do not. Weaver (1983), for example, reports that the average prospective education major ranked at about the 37th percentile of all SAT test-takers in 1972. Ballou and Podgursky (1997) find that the score of the average prospective education major fell at about the 45th percentile of all SAT-takers in 1980, but that this ranking improved considerably by 1992. While these comparisons are informative, they are also likely a noisy measure of practicing teacher quality, as not all SAT/ACT-takers attend or complete college, and those who do may change majors, or may never enter the teaching force. These studies also tend to neglect other potentially interesting aspects of the distribution of new teachers, beyond the mean.

Vance and Schlechty (1982) were the first to use a longitudinal study of college graduates to compare the academic ability of teachers and non-teachers. The use of longitudinal data permitted the authors to follow students into the workforce, avoiding the problems inherent in the use of SAT scores of intended education majors. Their tabulations on college graduates from the National Longitudinal Study of the High School Class of 1972 (NLS-72) indicate that teachers identified in 1979—particularly those who had expressed an intention to continue teaching—came disproportionately from the bottom two quintiles of the SAT score distribution.[9]

Several other studies have found a negative relationship between academic ability and the likelihood of entering teaching among cross-sections of college graduates. Manski (1985), in a test of how increased teacher pay might impact the quality of the teaching force, found a negative, statistically significant relationship between SAT/ACT scores and entry into teaching among working college graduates in the NLS-72. Hanushek and Pace (1995) and Vegas, Murnane, and Willett (2001) obtain similar results among college graduates using data from High School and Beyond. Podgursky, Monroe, and Watson (in press) find the same relationship in Missouri state administrative data.

The picture emerging from this cross-sectional literature is that the test scores of students choosing teaching as a profession during the 1970s and 1980s did not compare favorably to those of their college graduate peers. In light of the continuing trend of occupational desegregation described above, however, we would like to know whether or not this relationship between academic ability and entry into teaching has in fact worsened over time, and how the overall composition of teachers may have been affected as a result. The literature on this question is limited, and all papers thus far have considered only cohorts of college graduates over a relatively short period of time.[10] Murnane et al. (1991) and Bacolod (2003) are among the few to address the question of changes in teacher quality over time. Each use college graduates from the National Longitudinal Studies of Young Men, Women,

[9] This finding, when contrasted with those of Weaver (1983) and Ballou and Podgursky (1997), suggest that mean SAT scores of education majors may overstate the math and verbal abilities of practicing teachers.
[10] Pavalko (1970), using the Wisconsin Longitudinal Study, is an exception. Dividing the high school class of 1957 into three ability groups, he finds that at that time in Wisconsin, teachers were drawn disproportionately from the higher third of the IQ distribution.

and Youth and find that the percentage of graduates (of any AFQT score) entering teaching fell over the period 1967–1989, with a greater decline among those with the highest AFQT scores. In another recent strand of literature, Lakdawalla (2002) and Stoddard (2003) interpret declines over time in the relative wages of teachers as declines in the quality of teachers. Of course, these latter findings depend on a willingness to accept relative wages as a measure of teacher quality—and, as some authors have pointed out, the relationship between teacher salaries and student outcomes is a tenuous one (see Ballou and Podgursky, 1997; Hanushek, Kain, and Rivkin, 1999; Manski, 1985).[11]

DATA AND METHODS

Our data consist of longitudinal surveys of five cohorts of high school graduates—the Wisconsin Longitudinal Study (WLS) for the class of 1957, Project Talent (Talent) for the classes of 1960 to 1964, the National Longitudinal Study of the High School Class of 1972 (NLS-72), the sophomore cohort of High School and Beyond (HSB) for the class of 1982, and the National Education Longitudinal Study of 1988 (NELS) for the class of 1992.[12] These five studies are alike in that they each include results from a questionnaire administered during high school (usually during the senior year), all require students to participate in a battery of aptitude tests, and all conduct numerous follow-up surveys after high school. The inclusion of standardized test scores for all students allows us to place graduates into a cohort skill distribution, and assess how the quality of new teachers has changed over time (we use both centile rankings and standardized “z” scores).

We approach this question in two closely related ways. First, we estimate how the propensity for graduates with high test scores to enter teaching has changed over time. We accomplish this through the estimation of a series of logit models, where we compute the likelihood that a graduate of each cohort enters the teaching profession, conditional on their relative test score ranking.[13] Second, we look directly at the skill distribution of new female teachers, to see how the average academic ability and the ability composition of young female teachers has changed over time. While our primary focus is on women, we also perform the same analysis on our sample of men.

In this exercise, one goal is to hold as constant as possible across years the pool of women available for teaching. Although a college degree is usually a prerequisite for a teaching position, using samples of college graduates would potentially bias our results since college entrance rates have risen dramatically over time. One would expect results to be further biased to the extent that college completion has risen differentially across the test score distribution. By our calculation, college completion was proportionately larger for students in the bottom half of the distribution—lower scoring females were more than five times as likely to enter and finish college in 2000 than in 1964, in comparison to high scoring females who were

[11] Loeb and Page (2000) argue that cross-sectional studies of the relationship between teacher salaries and student achievement that do not take into account alternative wage opportunities are misspecified. They find that once relative wages are controlled for, that teacher salaries are modestly related to student outcomes.
[12] More information about these longitudinal studies can be found in the appendix. The Department of Education surveys (NLS-72, HSB, and NELS) intentionally oversampled certain minority and socioeconomic groups. These datasets each provided sampling weights, which we use where appropriate.
[13] This gradient comparison approach is similar in style to Murnane et al. (1991), who study the interaction of graduation year and AFQT score and its relationship with the propensity to teach.

only twice as likely.[14] Assuming the gradient between entry into teaching and academic ability is negative among college graduates (as was suggested by the cross-sectional studies discussed in the last section), one might estimate this gradient to be less negative among college graduates in later years simply as a result of an influx of more low-ability students into the sample. Restriction of the sample to working college graduates could add yet another layer of this kind of sample selection bias, by the same reasoning.

These potential biases are avoided by taking as the sample those individuals (both working and non-working) with at least a high school diploma.[15] As noted, high school graduation rates have increased only modestly since 1957, at least in comparison with college graduates (from 68 percent to 86 percent in Table 1—an increase driven largely by gains in high school completion among African American women). Admittedly, any change in the ability composition of high school graduates could bias results in a manner similar to that described above. There is no way, unfortunately, to test the extent of change in the composition of high school graduates in our data, but we believe that any bias induced by such changes is smaller (particularly among white women) than that potentially induced by limiting our analysis to college graduates.

Table 2 provides some descriptive statistics for the five cohorts of female high school graduates in our sample. To allow for comparison across surveys and avoid life-cycle effects on occupational choice, we chose a follow-up survey conducted when most respondents were approximately twenty-six years of age. These surveys were conducted in 1964 (for WLS respondents), 1971–1974 (Talent), 1979 (NLS-72), 1992 (HSB), and 2000 (NELS). As mentioned above, our measure of individual academic ability is a centile ranking or standardized score based on the combined math and verbal portions of a standardized test administered in high school. Table 2 provides the mean centile rank and standardized score for the high school graduates in our sample.[16]

This collection of longitudinal data is not without its imperfections. For one, these surveys provide snapshots of only five cohorts over this period. In addition, our oldest dataset (WLS) consists exclusively of non-Hispanic whites in Wisconsin—clearly not a nationally representative sample of high school graduates.

As a test of the applicability of the WLS data, in Table 3 we contrasted the 1936–1942 birth cohort (approximately the same cohort as our WLS respondents) of Wisconsin-born females to non-Wisconsin–born white females in the 1970 (1 percent) and the 1980 and 1990 (5 percent) Census Public Use Micro Samples (PUMS). Nothing in this table suggests that white Wisconsin women of this cohort look markedly different from white women of the same birth cohort born outside of Wisconsin.[17] In the most relevant year (1970), when this cohort was approximately 30 years of age, little difference can be seen in the educational attainment, labor force

[14] Results available from the authors upon request.

[15] Vegas, Murnane, and Willett's (2001) is one of the few studies that begins with a sample of high school graduates. However, like most of the empirical work described in the previous section, they do not look at changes in teacher quality over time, but rather the impact of certain individual characteristics on various sequential decisions (high school completion, college graduation, and finally entry into teaching).

[16] Means differ slightly from 50 and zero due to varied response to the follow-up surveys (while centile ranks were determined using all female graduates for whom we have a test score, only those responding to the follow-up survey are counted in these statistics). While only follow-up participants with test scores were selected for our sample, few observations were lost with this restriction. Response rates were quite high for the follow-up surveys—the lowest rate of response was 86.5 percent (WLS).

[17] Keep in mind that large samples as those we have used in Table 3 will tend to find statistically significant differences, even when the difference itself is quite small.

Table 2. Descriptive statistics for five cohorts of female high school graduates.

	WLS	Talent	NLS-72	HSB	NELS
High school graduation year(s)	1957	1960–63	1972	1982	1992
Follow-up survey year(s)	1964	1971–74	1979	1992	2000
Sample size	4609	1634	6751	5389	4284
Average age in years at follow-up	25.0	26.8	26.1	27.5	26.1
Race/ethnicity (percentage):					
White	—	88.1	82.1	74.5	74.6
Black	—	6.4	9.6	12.5	10.5
Hispanic	—	1.3	3.2	10.9	10.6
Other	—	4.2	5.0	2.1	4.4
Fraction with at least 2 years of college	28.6	29.7	40.6	35.6	59.9
Fraction with at least 4 years of college	14.5	20.5	24.7	26.0	42.3
Fraction working full- or part-time	38.9	48.7	72.1	74.1	86.8
Mean, mother's education (years)	10.3	—	12.4	12.6	13.1
Mean, father's education (years)	9.7	—	12.8	13.0	13.6
Fraction currently married (HS graduates)	82.9	80.2	60.8	58.5	46.4
Fraction currently married (college graduates)	65.3	77.8	46.7	50.5	40.3
Mean, centile ranking					
High school graduates	50.0	50.4	50.4	50.7	50.5
College graduates	73.4	74.9	72.4	76.2	67.4
Mean, standard score					
High school graduates	0.019	0.000	0.000	0.010	0.000
College graduates	0.820	0.843	0.742	0.914	0.579

Sample consists of those individuals who (a) have completed high school by the given follow-up year, (b) responded to the follow-up, and (c) had a valid test score. Centile ranks are based on students' placement in the distribution of all graduates of the same gender in their high school cohort. Sampling weights were used in generating centile ranks and means in NLS-72, HSB, and NELS surveys.

participation, or fraction teaching between the two groups, with the exception being the fraction earning master's degrees, which is almost one percentage point higher among non-Wisconsin–born females. Earnings and wages are slightly higher among non-Wisconsin–born females (by about 7 to 8 percent) in 1970, likely explained in part by their marginally higher educational attainment and greater tendency to reside in urban areas. Wider differences between these cohorts appear later in life (1980 and 1990), particularly in the fraction completing higher degrees. While we are unable to claim WLS to be nationally representative, we can have some confidence that this sample—at least at the age when most decisions about entering the teaching profession are made—looks much like the larger population of white females with high school degrees. The inclusion of Project Talent (which is nationally representative and also pre-1965) should provide a meaningful comparison.

While the test or batteries of tests administered in each survey do differ from one another, they are all quite similar in content to standardized tests like the SAT and ACT. Indeed, among those students for whom we have both a test score and an SAT/ACT score, the correlation between these scores is quite high (0.84 to 0.86). As an additional test of the comparability of these scores across surveys, we used logistic regression analysis to estimate the relationship between these centile rankings (or standardized scores) and entry into medicine among men across four of the five surveys. With the fraction of male high school graduates eventually entering the medical profession remaining roughly constant since 1970, and having no reason to believe that the relationship between cognitive ability and entry into medicine has

Table 3. Comparison of white female high school graduates born in Wisconsin and white females born elsewhere, 1936–1942 birth cohorts.

	Females, ages 28–34 in 1970, born in:		Females, ages 38–44 in 1980, born in:		Females, ages 48–54 in 1990, born in:	
	WI	All other areas	WI	All other areas	WI	All other areas
Sample size	1,494	51,174	7,475	271,924	7,939	295,978
Percentage unemployed	3.85	3.75	3.55	3.73	2.74	3.14
Percentage out of the labor force	56.5	56.4	29.2	32.3*	22.8	26.8*
Percentage of high school graduates that are:						
Teachers	10.1	9.1	6.8	7.1	6.8	7.0
Elementary teachers	6.6	6.3	4.8	5.4*	5.6	5.5
Secondary teachers	3.2	2.5	1.4	1.3	0.6	0.9*
Fraction with 4-year college degree	16.5	16.7	16.8	18.5*	17.9	19.8*
Fraction with master's degree	1.4	2.2*	3.9	4.7*	5.95	8.0*
Real wage/salary income (1990 dollars)	15,262	16,463*	15,345	16,677	18,455	20,897*
	(8,509)	(8,647)	(9,919)	(10,367)	(13,884)	(15,008)
Real weekly wage (1990 dollars)	310.19	333.80*	307.54	334.64*	369.35	418.24*
	(174.90)	(175.52)	(198.49)	(207.48)	(277.02)	(300.42)
Fraction living in MSA	59.7	64.9*	69.2	69.4	56.4	63.5*
Fraction living in same state as birth	67.5	57.9*	65.6	56.4*	65.4	55.6*

Source: 1970, 1980, and 1990 PUMS. (*) Identifies a statistically significant difference in means. Person weights were used in the 1990 samples. The 1970 and 1980 samples were designed to be self-weighting. Mean wages and income were computed using those earning at least $1000/year, and working at least 40 weeks per year.

changed much since 1960, we would expect—if test scores measure similar aptitudes—to see a fairly consistent relationship between these variables over time.[18] As Table 4 shows, this is indeed the case. In all cases, the coefficient on test score is positive and statistically significant (evidence that the qualities measured by these scores are strongly associated with entry into a high-cognitive ability profession), and the marginal effect (also shown in elasticity form) of test score on entry into medicine among men remains roughly constant across the four surveys.

Finally, our definition of occupation used with these samples is a broad one. Teachers and non-teachers are identified in these datasets using a self-reported occupation, which is reported regardless of the individual's labor force status. This broad definition is useful in this context if women who expect to spend more time out of the labor force self-select into flexible occupations like teaching (suggested by Flyer and Rosen [1997], Polachek [1981], and others).[19] In this case, a random sample of working females would be more likely to exclude teachers than non-

[18] Using data from the March Current Population survey (1970–1996), the fraction of male high school graduates age 25 to 34 who report themselves to be physicians or surgeons has remained very close to 0.5–0.6 percent, with no apparent upward or downward trend.

[19] Most empirical work on teacher quality to date restricts analysis to working women only. A caveat associated with our use of this broad occupational definition is that our results on teachers apply to those women who identify themselves as teachers, not necessarily those employed as teachers at the time of the survey. If some low-scoring women report themselves as teachers and are not in the labor force because they could not find work as a teacher, for example, then our results—as a general statement about practicing teachers—would bias the average ability of teachers downward.

Table 4. Logit estimates of the probability of male high school graduates entering medicine.

Dependent variable	WLS Physician in 1975		NLS-72 Physician in 1986		HSB Physician in 1992	
Sample mean	0.011	0.011	0.023	0.023	0.008	0.008
Sample size	4,331	4,331	4,358	4,358	4,834	4,834
Centile rank	0.052		0.051		0.055	
	(0.008)		(0.006)		(0.010)	
Standardized score		1.374		1.473		1.330
		(0.198)		(0.167)		(0.198)
Marginal effect of score at:						
Mean	0.0002	0.0065	0.0004	0.0124	0.0002	0.0045
85th centile or +1 SD	0.0013	0.0249	0.0023	0.0512	0.0011	0.0169
Elasticity of score at:						
Mean	0.0521	1.3676	0.0506	1.4603	0.0546	1.3250
85th centile or +1 SD	0.0509	1.3487	0.0486	1.4197	0.0537	1.3125

Standard errors in parentheses.

teachers. To the extent that labor force participation among teachers is correlated with academic ability, this sampling approach would likely result in a selection bias.[20] Note that teachers identified in our data may be elementary or secondary teachers, and may work (or have worked) for public or private schools.

Table 5 reports the top 10 most frequently reported occupations by female high school and college graduates in the WLS (1964) and HSB (1992) surveys. As expected, the evidence shows significantly less concentration in the occupations held by women in later years, particularly among those with a college education. Whereas 49 percent of female college graduates in the WLS sample were teachers in 1964 (compare to 52 percent in Figure 1), only 11.8 percent were teachers in the 1992 HSB follow-up (compare to 15 percent in Figure 1). In 1992, female college graduates were more likely to be in management (14 percent) or clerical work (a broad category, 17 percent) than in teaching.

RESULTS

The Quality of New Female Teachers

Table 6 presents some descriptive statistics for the female teachers we identified in each cohort. Not surprisingly, the centile ranking of the average female teacher lies consistently above that of the average high school graduate (50 and 0, by definition). We find, however, that the average female teacher in our sample scored consistently below the average female college graduate throughout this period (compare to Table 2). In addition, the ranking of the average new female teacher fell about 3 percentage points over this period, from the 67th centile in WLS (or 69th if using Project Talent)

[20] In the WLS sample, almost 35 percent of the women identifying themselves as teachers were out of the labor force at the time of the 1964 follow-up survey. Project Talent is unique in that most women out of the labor force did *not* report an occupation. We identified 58 "housewives" in the Talent survey who held teaching certificates—an indication that they might be teaching, if they were in the labor force at the time of the survey. However for consistency with the other surveys, we did not count these individuals as teachers. Incidentally, 18 of the 58 housewives (31 percent) with teaching certificates ranked in the top *decile* of the test score distribution.

Table 5. Most frequently reported occupations, two cohorts of female high school and college graduates.

	WLS (1964)		HSB (1992)	
	A. Female High School Graduates			
Rank	Occupation	Percentage	Occupation	Percentage
1	None/housewife	27.0	Clerical	34.2
2	Clerical	15.9	Service worker	16.3
3	Stenographer/typist	12.8	Manager	11.0
4	Teacher	9.7	Sales/worker	5.3
5	Nurse/professional	4.6	None/housewife	3.7
6	Bookkeeper	2.9	Operative (non-transport)	3.6
7	Office machine operator	2.1	Teacher	3.5
8	Waitress	1.9	Craftsman	2.6
9	Beautician	1.9	Health technician	2.2
10	Medical/dental technician	1.8	Nurse, dietician, therapist	2.2
	B. Female College Graduates			
1	Teacher	48.8	Clerical	17.3
2	None/housewife	14.2	Manager	13.9
3	Student	3.9	Teacher	11.8
4	Social worker	3.4	Sales	8.9
5	Stenographer/typist	3.1	Service worker	6.7
6	Nurse, professional	3.1	Accountant	5.4
7	Medical/dental technician	2.5	Nurse, dietician, therapist	5.1
8	College professor	2.2	Computer specialist	4.2
9	Therapist/healer	2.1	Writer/artist	4.1
10	Musician/music teacher	1.5	Health technician	2.3

to the 64th in the NELS—a drop of 5.2 percent.[21] As centile rankings mask information about the tails of the distribution, we also computed the mean standardized scores for teachers in each cohort. Here, the downward trend in the mean among female teachers is starker—a fall from 0.60 standard deviations above the mean female high school graduate in 1964 (or 0.65 in Talent) to 0.46 in 2000, a drop of 23 percent.

While the modest drop in the relative math and verbal abilities of the average new female teacher is of interest, it would also be informative to know how entry into the teaching profession changed differentially across the ability distribution. To examine this, we estimated five logit models (one for each cohort) where the probability of becoming a teacher is assumed to be a function of an individual's test score, age, and race.[22] Following Murnane et al. (1991), we begin by allowing cen-

[21] A test of the hypothesis that the mean centile rank of WLS teachers is greater than that of the NELS teachers is accepted at the 95 percent level ($p = 0.029$); for the mean standardized score, $p = 0.014$ for the same test. When male and female teachers are combined, the mean centile rankings in each cohort are (in order): 65.3, 65.7, 65.1, 66.4, and 64.0. The mean standard score for the combined teachers in each cohort are: 0.541, 0.515, 0.504, 0.566, and 0.472.

[22] Our restriction to three covariates was due in part to data limitations. We argue, however, that the question of interest here is how the *total* relationship between entry into teaching and academic ability has changed over time—not the partial relationship. Others, such as Murnane et al. (1991), include cognitive ability as one of a number of regressors, capturing instead the partial relationship between teaching and ability. We control for age due to some variation (especially in Talent) in the respondents' age. Due to its homogenous composition, race dummies were not relevant in the WLS regressions. Also, there were too few ethnic minorities identified in Talent to include race dummies as covariates.

Table 6. Descriptive statistics, new female teachers.

	WLS	Talent	NLS-72	HSB	NELS
High school graduation year	1957	1960–63	1972	1982	1992
Follow-up survey year	1964	1971–74	1979	1992	2000
Number of teachers	369	99	431	219	302
Fraction of all female high school graduates	8.0	6.1	6.8	4.1	7.0
Fraction of all female college graduates	55.2	30.4	24.4	13.3	16.3
Race of teacher (percentage)					
White	—	89.9	86.5	80.2	83.8
Black	—	7.1	8.4	10.7	5.3
Hispanic	—	—	1.4	7.9	8.0
Other	—	3.0	3.7	1.1	2.8
Fraction with at least 4 years of college	79.1	98.0	96.2	78.2	92.7
Fraction working	64.9	100.0	88.2	78.0	94.5
Fraction currently married	66.7	73.9	55.4	60.2	50.1
Mean, mother's education (years)	11.6	—	13.2	13.7	14.0
Mean, father's education (years)	11.3	—	13.5	14.5	14.7
Ability measures:					
Mean centile rank	67.2	69.5	66.4	64.8	63.7
Mean standard score	0.60	0.65	0.55	0.50	0.46

Teachers identified using the following codes: WLS (#51: Teachers, NEC), Project Talent (#400-433: All elementary and secondary teachers), NLS-72 (#142-145: Elementary, kindergarten, secondary, and teachers NEC), and NELS (#23: Educator, K-12). HSB teachers were identified by visual inspection, using verbatim responses. Means and centile ranks incorporate sampling weights in NLS-72, HSB and NELS.

tile rank (or standard score) to enter linearly into the model; we then relax this assumption and implement decile dummies as our measure of academic ability.

Table 7 provides coefficient estimates, marginal effects, and elasticities from our initial linear specification. Panel A displays the results where centile rankings are used, while panel B provides results where standardized scores are used. Marginal effects and elasticities are computed for a white female, of average age, at the 85th centile, which is one standard deviation above the mean.[23]

As one would expect in a sample of graduates with a minimum of a high school degree, we find in Table 7 that test scores and entry into teaching have a positive, statistically significant relationship across all five cohorts. What is more interesting in these tables is the substantial weakening of this relationship between ability and entry into teaching since 1964. On both an absolute and elasticity basis, the effect of a one point increase in centile rank (or standardized score) on the probability that a female high school graduate enters teaching has fallen markedly over this period, with the decrease much more prominent when standardized scores are used as a measure of ability. As shown in panel A of Table 7, the elasticity of a one-centile increase in test score calculated at the 85th centile has dropped from 0.021 in 1964 (with a 95 percent confidence interval of [0.017, 0.024]), to 0.014 in 2000 (with a 95 percent confidence interval of [0.010, 0.018]), or about a third of its size.[24] These

[23] We use white females as our basis for comparison, to allow for comparison with the WLS sample. The results differ little when race is excluded as a covariate. Elasticities—or the percentage increase in the likelihood of choosing teaching as an occupation for a unit change in test score rank—are more comparable here than marginal effects, as the mean of the dependent variable falls somewhat over the five cohorts. We use elasticities to make all of our cross-cohort comparisons.
[24] This result differs little when the elasticity is computed at the mean.

Table 7. Occupational choice model for females with at least a high school degree.

Dependent variable	WLS Teacher in 1964	Talent Teacher in 1971	NLS-72 Teacher in 1979	HSB Teacher in 1992	NELS Teacher in 2000
Sample mean	0.080	0.061	0.068	0.041	0.070
Sample size	4,609	1,634	6,751	5,389	4,284
A. Centile Rank					
Coefficient	0.024	0.028	0.021	0.018	0.017
	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Marginal effect,	0.0029	0.0028	0.0020	0.0009	0.0014
at 85th centile	(0.0004)	(0.0006)	(0.0003)	(0.0002)	(0.0003)
Elasticity,	0.021	0.025	0.019	0.017	0.015
at 85th centile	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Log–likelihood	−1197.0	−347.3	−1505.0	−758.7	−1059.7
Pseudo R^2	0.060	0.070	0.047	0.029	0.030
B. Standardized score					
Coefficient	0.685	0.807	0.643	0.493	0.486
	(0.063)	(0.121)	(0.055)	(0.090)	(0.069)
Marginal effect,	0.073	0.072	0.055	0.022	0.045
at +1 SD above mean	(0.009)	(0.015)	(0.007)	(0.005)	(0.008)
Elasticity,	0.602	0.727	0.581	0.470	0.436
at +1 SD above mean	(0.053)	(0.106)	(0.049)	(0.085)	(0.061)
Log–likelihood	−1200.1	−347.5	−1504.4	−759.2	−999.0
Pseudo R^2	0.058	0.070	0.047	0.029	0.045

Standard errors in parentheses. Sampling weights were using with the NLS-72, HSB, and NELS samples (unweighted regression results did not differ substantially from the weighted results). Other covariates included in these regressions include age and race dummies. Marginal effects and elasticities calculated for a white female high school graduate of average age.

effects are again a bit stronger when using standard scores as our measure of academic ability. Panel B indicates that the elasticity of a one standard deviation increase in test score calculated at one standard deviation above the mean has dropped considerably—from 0.602 (with 95 percent confidence interval of [0.513, 0.719]) to 0.436 (95 percent confidence interval of [0.317, 0.534]), or about 0.22 percentage points. The point estimate using Project Talent of 0.73 shows an even greater decline than that suggested with the use of WLS. Given our samples we are unable to reject the hypothesis that these elasticities do not differ from 1964 to 2000 at the 95 percent level (we can at the 90 percent level). Nonetheless, we would argue that the sharp decline in this strength of this relationship between test scores and teaching as an occupational choice is particularly striking.

Results thus far suggest that the relationship between academic abilities and entry into teaching among female high school graduates has indeed weakened over time. Our stronger results based on standardized scores also suggests that individuals near the tails of the test score distribution may be particularly important in explaining this trend. For this reason, and to allow for the possibility that test scores may in fact be nonlinearly related to the likelihood of entering teaching, we estimate our logistic model again for each cohort using decile dummies as our measure of ability.

As marginal effects and elasticities like those in Table 7 are less meaningful under this specification, we have instead calculated the (average) predicted probability that a female in each decile chose teaching as a profession, for all five cohorts (Table 8). For females in most deciles the probability of teaching fell roughly in half from 1964 to 1992, with a modest rise between 1992 and 2000, movements that closely follow general trends in enrollment over these years. We find, however, much larger drops in this probability for females in the top three deciles from 1964 to 1992, or the top, 8th, and bottom deciles from 1964 to 2000 (the introduction of the NELS cohort changes our findings for the 9th decile somewhat). Columns 6 through 10 of Table 8 normalize these predicted probabilities by the overall fraction of each cohort that we identified as teachers. Values greater than one indicate that the probability that a female high school graduate from that decile becomes a teacher is higher than at the mean; values less than one indicate the opposite. Here these trends are clearer—women in the top decile are much less likely to become teachers, relative to the average, in later vs. earlier years. The opposite trend is true for deciles near the bottom of the distribution (with the lowest decile being a notable exception).

Figure 2 plots these predicted probabilities. This picture illustrates particularly well the nonlinear relationship between test score and entry into teaching, and the comparatively larger reduction in the predicted probability of becoming a teacher among those at the top of the ability distribution.

How have these differential rates of entry into teaching affected the overall composition of new female teachers? Figure 3 looks at the distribution of new female teachers across decile groups. Approximately 32 to 35 percent of teachers scored in the lowest six deciles of the test score distribution in 1964, but we observe this proportion growing monotonically across our high school cohorts after 1971 to 1974. By 2000, nearly 42 percent of the teachers in our sample fell in the lowest six deciles of the test score distribution. The fraction of teachers in the 7th, 8th, and 9th deciles fluctuated over these five cohorts with few discernible trends. It does appear, however, that the fraction of teachers who scored in the 7th (and possibly 9th) deciles of their high school class has been steadily rising over this period (although the jump in the 9th decile appears only in our last cohort). Aside from the rising proportion of new female teachers who scored in the lower six deciles of the test score distribution, the most interesting trend in Figure 3 is the steadily declining share of new female teachers who scored in the top decile of their high school cohort. While in 1964 over 20 percent of young female teachers fell in the top decile of their high school class (over 25 percent when considering Project Talent), this number had fallen to 11 percent by 2000.[25]

If women near the top of their high school classes are becoming increasingly less likely to enter teaching, what career paths are they pursuing? We tabulated the top 10 occupations reported by females ranking in the top decile of their class in four of our five cohorts and the results are striking, if not unexpected. Whereas close to 20 percent of females in the top decile in 1964 chose teaching as a profession (teaching was the most frequently reported occupation among this group in 1964), only 3.7 percent of top decile females were teaching in 1992. Top scoring women in our 1992 cohort were much more likely to be working as computer specialists (5.9

[25] It is important to emphasize that the changes in test score rankings of new teachers are changes in the relative academic ability of new teachers, i.e., relative to other high school graduates. It has been noted that although the relative ability of new female teachers has been falling over time, the absolute skill level of new female teachers may be rising, if parental education is an input into cognitive ability (observe in Table 6 the steady increase in average parental education for these teachers). We thank Darius Lakdawalla for being the first to point this out.

Table 8. Predicted probabilities of entering teaching as an occupation by decile, females with at least a high school degree.

	Predicted Probabilities					Predicted Probabilities as Proportion of the Sample Mean				
	WLS	Talent	NLS-72	HSB	NELS	WLS	Talent	NLS-72	HSB	NELS
Sample mean	0.080	0.061	0.068	0.041	0.070	0.080	0.061	0.068	0.041	0.070
Decile of test score	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
10th	0.169	0.147	0.096	0.057	0.079					
	(0.018)	(0.028)	(0.012)	(0.010)	(0.013)	2.11	2.41	1.41	1.39	1.13
9th	0.135	0.111	0.109	0.054	0.145					
	(0.016)	(0.025)	(0.012)	(0.009)	(0.016)	1.69	1.82	1.60	1.32	2.07
8th	0.122	0.092	0.117	0.046	0.069					
	(0.017)	(0.022)	(0.011)	(0.009)	(0.013)	1.53	1.51	1.72	1.12	0.99
7th	0.094	0.061	0.089	0.062	0.112					
	(0.012)	(0.018)	(0.011)	(0.010)	(0.016)	1.18	1.00	1.31	1.51	1.60
6th	0.090	0.049	0.079	0.041	0.089					
	(0.013)	(0.016)	(0.010)	(0.008)	(0.014)	1.13	0.80	1.16	1.00	1.27
5th	0.079	0.063	0.068	0.047	0.062					
	(0.015)	(0.018)	(0.009)	(0.008)	(0.012)	0.99	1.03	1.00	1.15	0.89
4th	0.045	0.037	0.048	0.024	0.071					
	(0.009)	(0.014)	(0.007)	(0.007)	(0.012)	0.56	0.61	0.71	0.59	1.01
3rd	0.021	0.012	0.029	0.021	0.049					
	(0.007)	(0.007)	(0.006)	(0.006)	(0.010)	0.26	0.20	0.43	0.51	0.70
2nd	0.024	0.018	0.018	0.022	0.019					
	(0.007)	(0.006)	(0.004)	(0.007)	(0.007)	0.30	0.30	0.27	0.54	0.27
1st	0.022	0.019	0.001	0.017	0.007					
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	0.28	0.31	0.02	0.42	0.10
Ratio of top to bottom:										
Half	3.28	3.14	2.97	2.44	2.35					
Quartile	6.24	7.09	6.01	4.06	4.45					
Quintile	6.58	7.09	5.89	4.20	5.90					
Decile (10/2)	7.04	8.17	5.33	2.59	4.16					

Standard errors in parentheses. Values in columns (1)–(5) are the average predicted probability of entering the teaching profession, conditional on test score decile, for a female with at least a high school degree. Sampling weights were used in the original logit regressions for the NLS-72, HSB, and NELS samples (unweighted results did not differ substantially from the weighted versions). Ratio of deciles in bottom panel is the ratio of the 10th to 2nd decile.

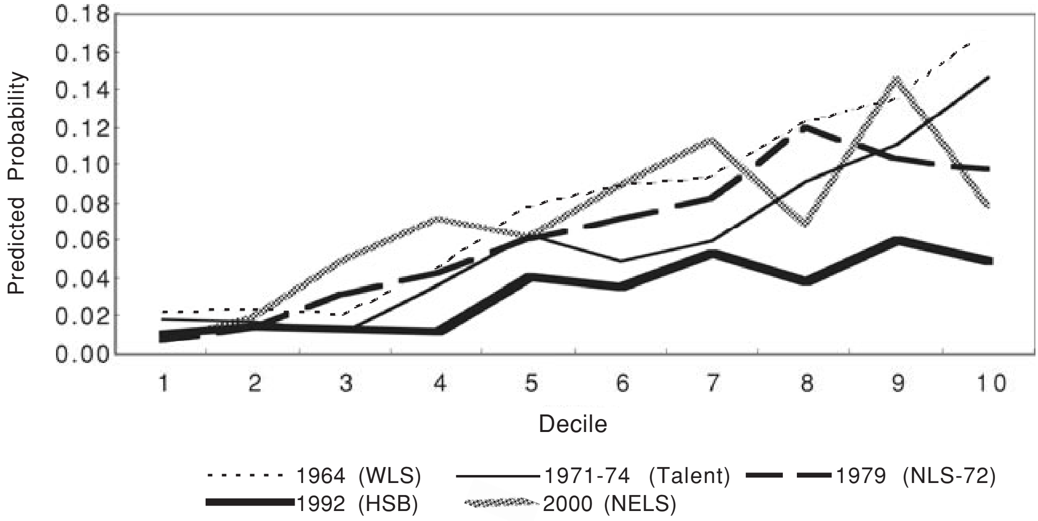


Figure 2. Predicted probability of becoming a teacher, by decile.

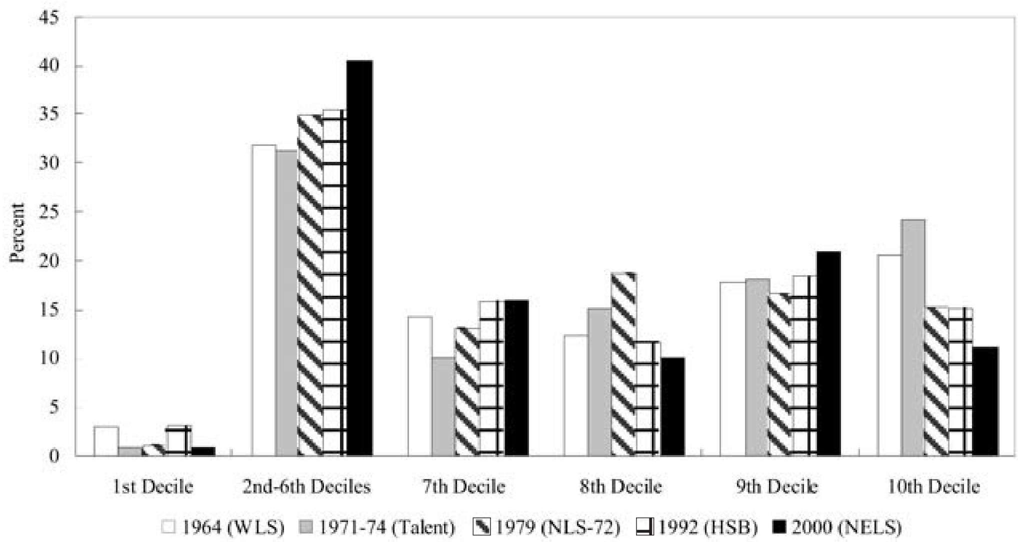


Figure 3. Distribution of teachers across decile groups, 1964–2000.

percent), accountants (6.0 percent), or managers (15.1 percent).[26] Top decile females were almost as likely to be lawyers (3.2 percent) as teachers.

THE QUALITY OF NEW MALE TEACHERS

An intriguing side effect of the gender desegregation of occupations and the movement of talented women into high-cognitive ability occupations is the potential substitution of high skilled men into teaching. As men lost a virtual monopoly on certain professions, finding themselves in competition with capable women for jobs, some may have opted for a career in teaching. While our sample sizes are obviously much smaller for male teachers, we summarize some of our findings on male teachers in Table 9.

The results in Table 9 are quite interesting, if only suggestive. Across these five cohorts, the average relative test score ranking of male teachers rose from 1964 to 2000 by 6.6 percent, or 28.2 percent if using standardized scores. This increase in the average ability of male teachers also appears to be driven by those at the top of the ability distribution, as the next panel in Table 9 indicates. While (as with women) the probability that a male high school graduate entered teaching fell over this period for most decile groups (again with a slight increase between 1992 and 2000), here the decline in probability is much less dramatic for those in the top decile. While most other decile groups saw a decline in the likelihood of entering

Table 9. Characteristics of new male teachers, 1964–2000.

	WLS	Talent	NLS-72	HSB	NELS
Sample size	4,379	1,527	6,444	4,984	3,855
Number of teachers	169	69	135	62	87
Mean centile ranking	60.6	59.6	60.9	71.6	64.6
Mean standard score	0.39	0.30	0.37	0.78	0.50
Average predicted probability of teaching					
1st–2nd deciles	0.013	0.016	0.009	0.002	0.007
3rd–4th deciles	0.042	0.033	0.013	0.006	0.020
5th–6th deciles	0.038	0.059	0.025	0.009	0.024
7th–8th deciles	0.038	0.049	0.025	0.009	0.025
9th–10th deciles	0.063	0.067	0.027	0.025	0.038
Average predicted probability of teaching as a proportion of the sample mean					
1st–2nd deciles	0.328	0.363	0.438	0.197	0.323
3rd–4th deciles	1.067	0.731	0.585	0.554	0.885
5th–6th deciles	0.976	1.306	1.182	0.898	1.070
7th–8th deciles	0.985	1.084	1.186	0.883	1.150
9th–10th deciles	1.624	1.485	1.247	2.490	1.733

Predicted probabilities are calculated in the same manner as in Table 8, and then averaged over the decile groups shown above. Sampling weights were used with the NLS-72, HSB, and NELS samples (unweighted results did not differ substantially from the weighted versions).

[26] Because we have not conditioned on labor force participation ("none" and "housewife" are counted as occupations), these tabulations should be unaffected by changes in female labor force participation. Interestingly—even in 1992—the most frequently reported occupation among top decile females was clerical work (17.2 percent).

teaching of 35 to 75 percent from 1964 to 2000, this reduction was only 29 percent for those in the top decile.

We calculated the net effect of these contrasting trends on the relative ability distribution of all teachers (females and males combined); the result was a somewhat dampened version of Figure 3. While the trends in each decile group share were qualitatively similar to those in Figure 3, the decline in the fraction of teachers coming from the top decile of the test score distribution was less sharp—when both male and female teachers are counted, top decile graduates comprised 18.9 percent of all teachers in 1964, and 13.7 percent of teachers in 2000. That the net effect is not larger is not surprising, given the overwhelming female share of the teaching profession.

Again, appropriate caution should be used with the results in Table 9. The sample of male teachers was naturally much smaller than that of females (as few as 62 male teachers in the case of HSB), and the representation of males (particularly top decile males) in the general population of teachers may be small enough that such a trend may not be economically significant. In addition, male teachers are much more likely to be secondary school teachers than elementary school teachers, suggesting that if anything, these results are most relevant for secondary level teachers. Nevertheless, we find our results on men to be intriguing, and worthy of additional study.

CONCLUSION

Despite a number of cross-sectional studies that have examined the characteristics of college graduates choosing to enter the teaching profession, there has been little empirical evidence on how these characteristics—particularly academic ability—have changed over a long period of time. We believe, in light of the vast occupational desegregation witnessed during the past four decades, that it is of great interest to understand how this desegregation may have affected the recruitment of highly skilled women into teaching.

In the results presented here we find some evidence of a slight but detectable decline in the relative ability of the average new female teacher, when ability is measured as one's centile rank in the distribution of high school graduates on a standardized test of verbal and mathematical aptitude. The magnitude of this decline is even greater when measuring ability using standardized scores. We also find that examination of the entire distribution of new teachers is more informative than trends in central tendency alone. Over the 1964–2000 period, women near the top of the test score distribution became much less likely to enter the teaching profession than their peers near the middle of the distribution. The apparent consequence has been a much lower representation of women of very high academic ability in the pool of elementary and secondary teachers. While the sample sizes of male teachers are much smaller, we detect the opposite trend among men.

If our results can be applied to the wider population of young teachers in the United States, a given student in 2000 (conditional on having a female teacher) could expect to find a teacher who is on average of only slightly lower academic ability than a given student in 1964. However, that student is much less likely to find a teacher of the highest academic ability than a student in 1964. Further, given recent research on the sorting of teachers across schools within states and school districts—the likelihood that a student in a low income or predominately black school encounters a teacher of the highest academic ability is likely to be even lower (See Lankford, Loeb, and Wyckoff, in press). For the casual observer, these results will surprise few. However, if the significant loss of women in the top decile—those who likely stood to benefit most from occupational desegregation—is indicative of a wider trend, then these findings should be of interest to parents, researchers, and policymakers alike.

This work was supported by grant SBR9811385 from the National Science Foundation, whose support the authors gratefully acknowledge. This paper also benefited from comments received in the "Teachers and Low Performing Schools" and "Teacher Quality" panels at the 2001 APPAM and 2004 ASSA annual meetings. The authors would also like to thank Julie Topieski, Seth Sanders, four anonymous referees, and members of the University of Maryland Labor and Public Finance seminar, for their helpful comments and suggestions.

SEAN P. CORCORAN is an Assistant Professor in the Department of Economics, California State University, Sacramento.

WILLIAM N. EVANS is a Professor in the Department of Economics, University of Maryland, College Park.

ROBERT M. SCHWAB is a Professor in the Department of Economics, University of Maryland, College Park.

REFERENCES

Bacolod, M. P. (2003). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality, 1960–2000. UCLA Department of Economics Working Paper, September.

Ballou, D., & Podgursky, M. (1997). Teacher pay and teacher quality. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Beller, A. (1992). Occupational segregation by sex: Determinants and changes. Journal of Human Resources, 17, 371–392.

Blau, F. D., Simpson, P., & Anderson, D. (1998). Continuing progress? Trends in occupational segregation over the 1970s and 1980s. Feminist Economics, 4, 29–71.

Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960s? Coleman Revisited. Economics of Education Review, 14, 1–21.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (ed.), Holding schools accountable: Performance-based reform in education (pp. 265–298). Washington, DC: Brookings Institution.

Flanagan, J. C., Tiedeman, D. V., Clemans, W. V., & Wise, L. L. (1981). Project TALENT public use file, 1960–1976. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Flyer, F., & Rosen, S. (1997). The new economics of teachers and education. Journal of Labor Economics, 15, S104–S139.

Goldin, C. (2004). From the valley to the summit: The quiet revolution that transformed women's work, Working Paper #10335. Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A. (1970). The production of education, teacher quality, and efficiency. In Do teachers make a difference? A report on recent research on pupil achievement (pp. 79–99). Washington, DC: U.S. Department of Health, Education, and Welfare.

Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement. American Economic Review, 61, 280–288.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. Journal of Economic Literature, 49(September), 1141–1177.

Hanushek, E. A. (1996). School resources and student performance. In G. B. Burtless (ed.), Does money matter? The effect of school resources on student achievement and adult success (pp. 43–73). Washington, DC: Brookings Institution.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). Teachers, schools, and academic achievement, Working Paper #6691 (revised). Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). Do higher salaries buy better teachers? Working Paper #7082. Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A., & Pace, R. R. (1995). Who chooses to teach (and why)? Economics of Education Review, 14, 101–117.

Hoxby, C. M., & Leigh, A. (in press). Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States. American Economic Review, Papers and Proceedings of the American Economic Association.

Krueger, A. (2002). Economic considerations and class size. Industrial Relations Section Working Paper #447. Princeton, NJ: Princeton University.

Lakdawalla, D. N. (2002). The declining relative quality of teachers, Working Paper #8263. Cambridge, MA: National Bureau of Economic Research.

Lankford, H., Loeb, S., & Wyckoff, J. (in press). Teacher sorting and the plight of urban schools: A descriptive analysis. Education Evaluation and Policy Analysis.

Lloyd, C. B., & Niemi, B. T. (1979). The economics of sex differentials. New York: Columbia University Press.

Loeb, S., & Page, M. (2000). Examining the link between teacher wages and student outcomes: The importance of alternative labor market outcomes. Review of Economics and Statistics, 82, 393–408

Manski, C. F. (1985). Academic ability, earnings, and the decision to become a teacher: Evidence from the national longitudinal study of the high school class of 1972, Working Paper #1539. Cambridge, MA: National Bureau of Economic Research.

Murnane, R. J., Singer, J. D., Willett, J. B., Kemple, J. J., & Olsen, R. J. (1991). Who will teach? Cambridge, MA: Harvard University Press.

Pavalko, R. M. (1970). Recruitment to teaching: Patterns of selection and retention. Sociology of Education, 43, 340–353.

Podgursky, M., Monroe, R., & Watson, D. (in press). Teacher mobility, pay, and academic quality. Economics of Education Review.

Polachek, S. W. (1981). Occupational self-selection: A human capital approach to sex differences in occupational structure. Review of Economics and Statistics, 63, 60–69.

Rice, J. K. (2003). Teacher quality: Understanding the effectiveness of teacher attributes. Washington, DC: Economic Policy Institute.

Rockoff, J. E. (in press). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, Papers and Proceedings of the American Economic Association.

Siltanen, J., Jarman, J., & Blackburn, R. M. (1995). Gender inequality in the labour market: Occupational concentration and segregation. Geneva: International Labour Office.

Stoddard, C. (2003). Why has the number of teachers per student risen while teacher quality has declined? Journal of Urban Economics, 53, 458–481.

Temin, P. (2002). Teacher quality and the future of America. Eastern Economics Journal, 28, 285–300.

Vance, V. S., & Schlechty, P. C. (1982). The distribution of academic ability in the teaching force: Policy implications. Phi Delta Kappan, 64(September), 22–27.

Vegas, E., Murnane, R. J., & Willett, J. B. (2001). From high school to teaching: Many steps, who makes it? Teachers College Record, 103(June), 427–449.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. Review of Education Research, 73(Spring), 89–122.

Weaver, W. T. (1983). America's teacher quality problem: Alternatives for reform. New York: Praeger Publishers.

APPENDIX
Data Sources

1. Wisconsin Longitudinal Study: available at <http://dpls.dacc.wisc.edu/WLS/>.
2. Project Talent Public Use File, 1960-1976: available through the Inter-University Consortium for Political and Social Research at <http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/07823.xml>. For more information about this unique dataset, see Flanagan et al. (1981).
3. National Longitudinal Study of the High School Class of 1972: available from the National Center for Education Statistics at <http://nces.ed.gov/surveys/nls72/>.
4. High School and Beyond Sophomore Cohort (Fourth Follow-up): available through the National Center for Education Statistics at <http://nces.ed.gov/surveys/hsb/>. We used the restricted use version of this data; however, all of the variables used in this analysis are available through the public use version.
5. National Education Longitudinal Study of 1988: available from the National Center for Education Statistics at <http://nces.ed.gov/surveys/nels88/>. We used the restricted use version of this data in this analysis.
6. Current Population Survey, 1964–2000: compiled by the Bureau of Labor Statistics. For more information, see <http://www.bls.census.gov/cps/cpsmain.htm>.
7. Integrated Public Use Microdata Series (IPUMS), 1980 and 1990 5 percent state samples (a 1-in-20 sample of the U.S. population) and the 1970 1 percent state samples (a 1-in-100 sample of the population). Information about this data is available at <http://www.ipums.umn.edu/usa/index.html>.