

NBER WORKING PAPER SERIES

CHANGING LABOR MARKET OPPORTUNITIES FOR WOMEN  
AND THE QUALITY OF TEACHERS 1957-1992

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

Working Paper 9180  
<http://www.nber.org/papers/w9180>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2002

This work has been supported by a grant from the National Science Foundation. The authors wish to thank participants of the “Teachers and Low Performing Schools” panel at the November, 2001 APPAM Conference in Washington, D.C., as well as those participants of the University of Maryland Labor and Public Finance seminar, for their helpful comments and suggestions. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Sean P. Corcoran, William N. Evans, and Robert S. Schwab. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Changing Labor Market Opportunities for Women and the Quality of Teachers 1957-1992  
Sean P. Corcoran, William N. Evans and Robert S. Schwab  
NBER Working Paper No. 9180  
September 2002  
JEL No. I20, J24, N32

**ABSTRACT**

School officials and policy makers 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 remained a predominately female profession for years; at the same time, the employment opportunities for talented women outside of teaching have soared. In this paper, we combine data from four longitudinal surveys of high school graduates spanning the years 1957-1992 to examine how the propensity for talented women to enter teaching has changed over time. We find that 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 from 1964 to 1992—by our estimation, from almost 20% to under 4%.

Sean P. Corcoran  
Department of Economics  
University of Maryland  
College Park, MD 20742  
corcoran@econ.umd.edu

William N. Evans  
Department of Economics  
University of Maryland  
College Park, MD 20742  
and NBER  
evans@econ.umd.edu

Robert S. Schwab  
Department of Economics  
University of Maryland  
College Park, MD 20742  
schwab@econ.umd.edu

*“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, August 17, 2001).*

## **I. 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 to teaching, and the debate over how best to attract and retain talented, better-qualified teachers seems to intensify. One explanation for this mounting frustration points to the remarkable gender desegregation of the labor market since 1960.<sup>1</sup> Schools that once found in women a captive labor pool are now forced to compete with more lucrative professions, with the best and brightest thought least likely to remain in teaching. This conjecture has been the impetus behind many suggested policy measures—salary increases, or relaxed testing and course requirements, for example—all 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 certain teacher characteristics—like academic ability—are quite important for student achievement.<sup>2</sup> This evidence is particularly remarkable given the current lack of agreement over the importance of other inputs into education, such as per-pupil expenditure and class size.<sup>3</sup> Recognizing the apparent importance of attracting quality teachers, the federal

---

<sup>1</sup> Temin (2002), for example, 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.

<sup>2</sup> See, for example, Hanushek (1970, 1971), Ehrenberg and Brewer (1994, 1995), Ferguson and Ladd (1996), and Hanushek, Kain, and Rivkin (2000). Wayne and Youngs (2001) and Goldhaber (2002) review recent literature on the relationship between teacher characteristics and student achievement.

<sup>3</sup> Hanushek (1986, 1996) and Krueger (2002) provide contrasting views of this issue.

government responded in 2001 with Title II of the Elementary and Secondary Education Act—a \$3.175 billion program specifically targeted toward improving teacher quality.

The hypothesis that gender desegregation of the professions has reduced the number of high-ability women choosing to teach is so frequently cited that it has virtually earned the status of common knowledge. However, there is surprisingly little evidence to date on the extent to which occupational desegregation has in fact affected teacher quality over time. This lack of evidence was noted in the conclusion to a recent paper by Podgursky, Monroe, and Watson (2001), who state:

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 teaching profession. Unfortunately, time-series data on teacher quality are not available to directly test this hypothesis.

In this paper, we combine data from four longitudinal surveys of high school students spanning almost four decades (1957-1992), to see how the propensity for high test scoring women to teach has changed over time. Inclusion of data prior to 1970 allows us the opportunity to provide evidence as to how this relationship between academic ability and entry into teaching has changed over a period of vast gender desegregation of occupations.<sup>4</sup>

Our results show that examination of the entire distribution of teacher quality and its changes over time tells a much richer story than one could tell from measures of central tendency alone. While we detect only a slight decline in the *average* ability of female entrants into teaching over this period, the likelihood that a female at the *top* of the ability distribution—presumably the type of individual most likely to benefit from new opportunities in the labor

---

<sup>4</sup> 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, Bacolad (2001) relates changes in relative teacher salaries to changes in the quantity and quality of teachers over time.

market—has plummeted, with a female in the top decile of her high school class less than one-fifth as likely to enter teaching in 1992 versus 1964.<sup>5</sup> Whereas 20 percent of all females in the top decile of their high school class went into teaching in 1964, only 3.7 percent of this group chose to teach in 1992.

We organize this paper as follows. In section two, we begin with a discussion of changes in the labor market since 1960 and look at the way these changes may have affected the quality of incoming teachers. Section three describes some of the existing research on teacher quality, discusses how other researchers have approached the rather nebulous measure of teacher “quality,” and explains how our approach will differ from that of previous research. Section four describes our methodology and data. Section five presents our results of a simple econometric model relating the academic ability of females to entry into teaching, and provides a brief look at the male teachers in our data. In section six, we make some concluding remarks.

## **II. Background**

As is widely known, women’s relationship with the labor market fundamentally changed during the forty-year period considered here. Labor force participation among women aged 25-34 nearly doubled between 1964 and 1992; at the same time the fraction of women 25-34 with (four-year) college degrees tripled (see Table 1). By comparison, the fraction of college educated young men increased by only 50 percent over the same period. These trends were

---

<sup>5</sup> The fraction of high school graduates who eventually enter teaching has fallen for almost all ability levels. However, as a comparison, females in the middle of the distribution (4<sup>th</sup>-7<sup>th</sup> centiles) were only about half as likely to enter teaching.

likely aided in part by key legislative movements of the early 1960's and 70's that altered significantly the landscape of occupational opportunity for women.<sup>6</sup>

Occupations became considerably less gender segregated over this period.<sup>7</sup> Table 1 shows an index of gender representation in two professions—medicine and law—calculated as the percent female in the occupation divided by the percent female of the entire labor force, using data on men and women aged 25-34 in the March Current Population Survey (CPS), 1968-1996. A value less than one indicates that women are *underrepresented* in the occupation (within this age range), relative to their representation in the labor force.<sup>8</sup> In the late 1960's this index ranged from 0.20 to 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 (0.79 and 0.94, respectively).

Despite striking shifts in the gender composition of professions such as medicine and law, the gender composition of new teachers has remained roughly constant. Females continue to dominate the teaching profession, comprising roughly 70 percent of teachers aged 25-34—a fraction virtually unchanged since 1964 (Table 1). The proportion is higher for elementary teachers (82-85 percent) than secondary teachers (46-57 percent), though the fraction has been steadily falling for elementary and rising for secondary teachers within this age range.

While teaching remains a predominantly female occupation, the occupation itself has significantly diminished in importance as a career path for female college graduates. Nowhere is this so evident as in the plummeting percentage of (working) female graduates who identify

---

<sup>6</sup> 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.

<sup>7</sup> This is well documented in Blau, Simpson, and Anderson (1998), Beller (1992), and elsewhere.

<sup>8</sup> This and other methods of measuring occupational segregation and concentration are discussed in Siltanen, et. al. (1995).

themselves as teachers (Figure 1). According to data from the March CPS, in 1964 over *half* of working female college graduates were teachers—by 1996, this percentage had fallen below 15 percent. While this drop 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 1960's, most went into teaching; of those completing college today, most do not. This raises the question—how has the composition of those women who *do* choose to teach changed over time?

Occupational desegregation and the movement of females into high-skill professions such as medicine and law do not necessarily imply a reduction in the pool of talent available for teaching, if the pool of college graduates is expanding. While Table 1 shows a dramatic increase in the pool of female college graduates, the growth in teachers over the same period was not nearly as fast. The total stock of teachers rose by only 89 percent between 1960 and 1996, rising the most between 1960 and 1970 (46 percent) to accommodate an enrollment boom, and rising slowly thereafter. Meanwhile, the fraction of young women with degrees tripled, with the most talented more likely than ever to earn a degree. Figure 2 shows college completion rates among four cohorts of female high school graduates, conditional on a high school standardized test score decile.<sup>9</sup> While in 1964 only 23-42 percent of women in the top three deciles completed college, by 1992 47-72 percent of these women were finishing college. It remains to be seen whether this influx of talented females allowed the teaching profession to maintain a certain level of quality, or whether these new graduates increasingly selected into occupations other than teaching.

---

<sup>9</sup> The data used in Figure 2 will be described in greater detail in section four. For now, this figure is presented solely for illustrative purposes.

### **III. Prior Literature on Teacher Quality**

Concern over teacher quality is not new. Weaver (1983) recounts a nearly 100-year history in the United States of attempts to raise teacher standards. Much of the research on teacher quality in recent years has focused on 1) what “quality” exactly refers to, and how it should be measured, 2) how those entering teaching compare along various dimensions to their non-teacher peers, 3) whether teacher characteristics are in fact important for student achievement, and 4) how best to raise the quality of the teaching force without drastically reducing the supply of teachers or the diversity of the profession.

Ideally, teacher quality would be measured as a multi-dimensional vector of all teacher characteristics that are positively associated with outputs of the educational process. This vector likely includes many attributes—like patience, creativity, or communication skills, for example—that are unobservable or otherwise difficult to measure. Recognizing the shortcomings of the approach, researchers have focused instead on such measurable characteristics as degree attainment, certification status, teaching experience, school selectivity and test scores in this vector of teacher “quality.” Others, like Lakdawalla (2001) and Stoddard (2001) make use of relative teacher wages as proxies for teacher quality.

In this paper, we use a measure of academic ability—scores from a standardized test administered in high school—as our measure of “quality.” We acknowledge the obvious limitations of using a one-dimensional measure of quality—to think that a single test score can capture all of the factors that make for an effective teacher would be naïve at best—but it is difficult to believe that measures, say, 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. And there is a growing literature that finds that such measures—particularly teachers’ verbal ability scores—are among the most important determinants of student achievement.<sup>10</sup> 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 roughly 4 to 8.5 percent. Likewise, Ferguson and Ladd (1996) find a positive relationship between Alabama teachers’ ACT composite scores and student reading score gains from 3<sup>rd</sup> to 4<sup>th</sup> grade. What a teachers’ 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.

Discovery of the apparent importance of teachers’ academic ability in the education production function comes at the same time a number of other studies show the academic ability of teachers and aspiring teachers did not compare favorably to that of their peers in the 1970’s and 80’s. These conclusions have often been made through comparison of the mean SAT or ACT scores of examinees intending to major in education to those who do not. Weaver (1983), for example, reports that the average prospective education major ranked at about the 37<sup>th</sup> percentile of all SAT verbal test takers in 1972, a ranking that remained unchanged through 1980. Ballou and Podgursky (1997) find that the average SAT score of prospective education majors fell at about the 45<sup>th</sup> percentile, but that this ranking improved considerably by 1992. While these comparisons are informative, they may be misleading as not all SAT/ACT takers attend college, and those who do may change majors, or may never enter the teaching force.

---

<sup>10</sup> See, for example, Hanushek (1970, 1971), Ferguson and Ladd (1996), Ferguson (1998), Ehrenberg and Brewer (1994, 1995), Hanushek, Kain, and Rivkin (2000), Wayne and Youngs (2001) and Goldhaber (2002).

These studies also tend to neglect other potentially interesting aspects of the distribution of teacher quality, other than the mean.

Vance and Schlechty (1982) were one of the first to make use of a longitudinal study to compare the academic ability of teachers and non-teachers. The availability of longitudinal data permitted the authors to actually follow students into the workforce, avoiding the problems inherent in using SAT scores of intended education majors.<sup>11</sup> Their tabulations on college graduates from the National Longitudinal Study of the High School Class of 1972 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.

Several other cross-sectional studies have found a negative relationship between academic ability and the likelihood of entering teaching among college graduates. Manski (1987), in a test of how increased teacher salaries 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 (2001) find the same relationship in Missouri state administrative data.

The picture emerging from this literature is that the test scores of students choosing teaching as a profession during the 1970's and 80's did not compare favorably to those of their college graduate peers. In light of the continuing trend of occupational desegregation described in section two, however, we would like to know whether or not this relationship between academic ability and entry into teaching has in fact *worsened* over time. The literature on this

---

<sup>11</sup> Gitomer and Latham (1999) are able to link SAT and ACT scores to practicing teachers by linking test score data to prospective teachers who elected to take the Praxis I or II licensing exam. Unfortunately, this data is only available for those who took the licensing exam between 1994 and 1997.

question is quite limited, and virtually no research has included data on graduates much before 1979.<sup>12</sup> Murnane, et. al. (1991) and Bacolad (2001) 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, and Youth and find that the percentage of graduates (of any IQ) entering teaching fell over the period 1967-1989, with a greater decline among those with high IQ scores. Lakdawalla (2001) finds a decline in the schooling and human capital value of teachers, relative to other workers, among the 1910-1950 birth cohorts of the U.S. Census. He interprets this as a decline in the quality of teachers, brought about through the substitution of the quantity of teachers for quality teachers over time, as the productivity of teachers' general knowledge declined relative to the specialized knowledge of other college graduates. Stoddard (2001) interprets a similar decline in the wage rank of teacher salaries as a decline in teacher quality. Of course, these findings depend on a willingness to accept relative wages as a measure of teacher quality.

Here, we hope to improve on the existing literature by bringing together for comparison many of the datasets mentioned above (and incorporating some older ones), measuring teacher quality with a variable believed to have important effects on student achievement, and by taking care to keep the underlying sample constant over time. Our methodology and data will be described in further detail in section four.

## **IV. Methodology and Data**

### Methodology

---

<sup>12</sup> 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.

To provide a picture of the changing relationship between academic ability and entry into teaching, we combine data from four longitudinal studies, two of which are comprised of individuals who graduated from high school prior to 1965. Later in this section, we will look for changes across these four cohorts in the average ability and the ability distribution of those persons who identify themselves as teachers. Further, in section five we estimate—for each cohort—a simple econometric model of the likelihood that graduates enter teaching, conditional on their measured ability:

$$(1) \quad \Pr(Y_{it} = 1) = f(\alpha_t + \beta_t x_{it} + \gamma_t' z_{it}),$$

where  $Y_{it}$  is an indicator equal to one if person  $i$  (in cohort  $t$ ) identifies herself as a teacher,  $x_{it}$  is a measure of person  $i$ 's academic ability,  $z_{it}$  is a vector of student and/or parental characteristics, and  $f$  is the logistic function. Loosely, we are interested in whether or not there have been changes in the gradient  $\partial f / \partial x$ —a measure of the relationship between ability and entry into teaching in a given year—that would have affected the distribution of ability among new teachers.

In each cohort, we include those individuals with at least a *high school* diploma as our sample.<sup>13</sup> As we described in the introduction, college completion among females tripled from 1964-1996. Were we to use college graduates as a reference group, our interpretation of changes in the gradient  $\partial f / \partial x$  could be biased if college completion has increased differentially across ability groups. A look at Figure 2 suggests that increases in college completion rates were not constant across ability levels. While the absolute increase in the fraction of females completing college was greater for high scoring females, the increase was proportionately larger for students

---

<sup>13</sup> Vegas, Murnane, and Willett (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 section three, it does 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).

at the bottom of the distribution—low scoring females were four times more likely to complete college in 1992 vs. 1964, compared to twice as likely for high scoring females. Assuming that  $\partial f / \partial x$  is negative among college graduates (as was suggested by several cross-sectional studies discussed in section three), one might estimate  $\partial f / \partial x$  to be *less negative* among college graduates in later years simply as a result of the influx of more low-ability students into the sample. Similarly, one could overestimate a change in  $\partial f / \partial x$  if individuals with higher test scores were more represented in later year samples. Restriction of the sample to *working* college graduates could add yet another layer of this kind of sample selection bias, by the same reasoning.

We avoid these potential problems by using a sample of all (working and non-working) high school graduates—a sample we believe to be much more stable over this period. While the high school graduation rate among women also rose (from 68 percent to 86 percent, Table 1—an increase driven largely by gains in high school completion among black women), the growth was not nearly as dramatic. Admittedly, any change in the ability composition of high school graduates could bias our results in a similar manner. Unfortunately, there is no way using our data to detect the extent of change in the composition of high school graduates, but we believe that any bias induced by such changes is smaller (especially among white women) than that potentially induced by limiting our analysis to college graduates.

## Data

Our data consists of four separate cohorts of high school graduates from four longitudinal surveys—the Wisconsin Longitudinal Study (WLS) for the class of 1957, Project Talent (Talent) for the classes of 1960-64, the National Longitudinal Study of the High School Class of 1972

(NLS-72), and the sophomore cohort of High School and Beyond (HSB) for the class of 1982. These four studies are alike in that they each include a detailed survey of the student during their senior (or sophomore) year, all require students to participate in a battery of aptitude tests, and all conduct numerous follow-up surveys after high school. Together, these surveys provide us four distinct pictures of the relationship between academic ability (as measured in high school) and occupational choice among high school graduates over four decades.<sup>14</sup>

In each survey, we take as our sample women who have graduated from high school, have a test score available, and responded to a selected follow-up survey. To allow for comparison across surveys and avoid life-cycle effects on occupation choice, we choose the follow-up survey conducted when most respondents were approximately twenty-six. These surveys were conducted in 1964 (for WLS respondents), 1971-74 (Talent), 1979 (NLS-72), and 1992 (HSB). Descriptive statistics for each sample are provided in Table 2.<sup>15</sup>

Neither of our 1960's-era datasets is ideal. WLS, consisting exclusively of white high school graduates in Wisconsin is clearly not a nationally representative sample, but it is large—4,609 women participated in the 1964 follow up. By contrast, Project Talent is designed to be nationally representative, but publicly released data contains one-third as many observations as WLS. As a test of the comparability of the WLS data, we contrasted the 1936-42 birth cohort (approximately the same cohort as our WLS respondents) of Wisconsin-born females to non-Wisconsin born white females in the 1970 1% and the 1980 and 1990 5% Census PUMS (see Table 3). Nothing in Table 3 suggests that white Wisconsin women of this cohort look markedly

---

<sup>14</sup> More information about these longitudinal studies can be found in the Data Appendix. The Department of Education surveys (NLS-72 and HSB) intentionally oversampled certain minority and socioeconomic groups. These datasets each provided sampling weights, which we use where appropriate.

<sup>15</sup> Not many observations were lost in restricting the sample to follow-up participants with test scores. The lowest rate of response among female high school graduates (with test scores available) across the four surveys was 86.5 percent (WLS)—see Table 2.

different from white women of the same birth cohort born outside of Wisconsin.<sup>16</sup> In the most relevant year (1970), when this cohort was approximately thirty years of age, we see little difference in the educational attainment, labor force participation, or fraction teaching between the two groups, with the exception being the fraction earning masters' 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 MSAs. 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 a nationally representative sample, we can have some confidence that our 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. Together with Project Talent, we hope to have a fairly representative picture of female high school graduates in the late 1950's and early 60's.

Among other variables, we have for each student a measure of academic ability—relative to their high school peers—and a self-reported occupation at approximately age twenty-six. Exams were administered in high school to most or all participants of the original studies. As our measure of cognitive or academic ability, we took each student's combined raw score on math and verbal portions of these exams and computed a centile ranking and standard ( $z$ ) score based on their placement in the distribution of all female high school graduates. Table 2

---

<sup>16</sup> 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.

provides the mean centile rank and standard score of graduates responding to each follow-up survey.<sup>17</sup>

While the test or battery of tests administered in each survey do differ from each other, most—like the SAT or ACT—are tests of verbal and mathematical aptitude. Indeed, among those students for whom we have both a test score and an SAT or ACT score, the correlation between these scores is high (0.853 for the NLS SAT scores, and 0.823 and 0.828 for the NLS and HSB ACT scores, respectively). As an additional, admittedly rough test of the comparability of these test scores across surveys, we used logistic regression analysis to examine the relationship between centile rankings or z-scores and entry into medicine among men across the four surveys. With the fraction of male high school graduates eventually entering medicine remaining roughly constant since 1970 and having no reason to believe that the relationship between cognitive ability and entry into the medical profession has changed much since 1960, we would expect (if our test scores are measuring similar aptitudes) to see a fairly consistent relationship between these two variables over time.<sup>18</sup> And 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 such as medicine), and the marginal effect (also shown in percentage, or elasticity, terms) of test score on entry into medicine among men remains roughly constant across the four surveys.

---

<sup>17</sup> Means differ slightly from fifty and zero due to varied response to the follow-up surveys. Note that centile ranks were determined using *all* female high school graduates for whom we have a test score available, not just those who responded to the follow-up survey.

<sup>18</sup> Using data from the March Current Population survey, the fraction of male high school graduates age 25-34 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, over the period 1970-1996.



Our occupation variable has a fairly broad definition here and should be interpreted as “profession” or “line of work,” as the individuals in our sample report occupations, regardless of whether or not they are currently working. A broader definition is actually quite useful in this context, as it has been suggested by Polachek (1981) and Flyer and Rosen (1997) that women who expect to spend more time out of the labor force may select into occupations like teaching because of their flexibility.<sup>19</sup> If this is indeed the case, comparing teachers to non-teachers in a sample of *working* women may lead to biased results.<sup>20</sup> Note that teachers identified in our data may be elementary or secondary teachers, and may work (or have worked) for public or private school districts. Unfortunately, our data does not allow us to make an accurate distinction between these classes of teachers.

Tables 5a-b report the top ten most frequently reported occupations by female high school and college graduates in each of the four surveys. As expected, we find evidence here of significantly less concentration in the occupations held by women in 1992 vs. 1964, 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.

---

<sup>19</sup> In our WLS sample, virtually *half* of the women identifying themselves as teachers were out of the labor force at the time of the 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 Project 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 were ranked in the top *decile* of the test score distribution.

<sup>20</sup> Most empirical work on teacher quality to date restricts analysis to working women only. One caveat associated with our use of a broader definition of occupation is that it suggests 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 6 presents some preliminary findings on the female teachers we identified in our samples. While the mean centile rank and standard score lies consistently above that of the average high school graduate (fifty and zero, by definition), we see the centile rank of the average new female teacher falling about 3 percentage points, from the 67<sup>th</sup> to the 64<sup>th</sup> percentile—about 4.3 percent—over this time period.<sup>21</sup> As centile rankings mask information about the tails of the distribution, we also compared the mean standard scores of teachers across these cohorts. Here, the downward trend in the average among female teachers is more stark—a fall from 0.60 standard deviations above the mean (or 0.69 in Talent) to 0.48 from 1964 to 1992—conservatively, a 20 percent drop.

Figure 3 takes a closer look at the distribution of these teachers across ability groups. Over the four cohorts, approximately 35 percent of teachers scored in the lower six deciles of the test score distribution, a proportion relatively unchanged over time. Teachers in the eighth and ninth deciles each make up about 17-18 percent of teachers in all four years. The most visible change in Figure 3, however, is in the top decile, where in the 1960's 20-25 percent of elementary and secondary teachers scored. By 1992, only 10 percent of the female teachers in our data scored in this decile.<sup>22</sup>

Our results so far suggest that—while the academic ability of the average new female teacher has remained relatively unchanged over this period relative to the average high school graduate (when comparing mean centile ranks)—teachers are much less likely to come from the

---

<sup>21</sup> The magnitude of this decline in the centile rank of the average teacher may in fact be underestimated, if the ability composition of *high school* graduates has changed over time (and if the increase in the high school graduation rate is driven by students at the lower end of the distribution). But again, we are unable to test this hypothesis in this data.

<sup>22</sup> Keep in mind that this is the top decile of the *high school graduate* distribution—a sizable population, not a small ‘elite’ as the top decile of the college graduate distribution might be considered.

top decile of their high school class in 1992 vs. 1964. We investigate these findings in further detail in section five.

## **V. Results—New Entrants into Teaching**

### Female Entrants

In the last section, we found that the academic ability of the *average* new female teacher—measured as one’s placement in the distribution of female high school graduates—has fallen somewhat since 1964. Here, we find that this decline can be largely explained by a marked decrease in the likelihood of teaching among women at the highest end of the test score distribution—presumably those who stood to benefit most from changes in occupational opportunity over this period.

To investigate how the relationship between academic ability and entry into teaching has changed over time, we estimated equation (1) separately for each cohort, including at first a continuous measure of academic ability (the student’s centile ranking or standardized score), race dummies, and the individual’s parents’ years of education, where available.<sup>23</sup> Including controls for parental education allows us to control for the possibility that differential rates of entry into teaching are better explained by college opportunities afforded children of more educated or wealthier parents. We included other available covariates in this specification—such as family socioeconomic status, age, size of the community where the individual’s high school was located, whether or not the student’s mother or father were teachers, etc. (not shown in our

---

<sup>23</sup> Parental education was not provided in Project Talent. Talent did identify the respondent’s race, but there were too few ethnic minorities identifying themselves as teachers to include race dummies in this regression. In regressions where race dummies are included, whites are the omitted group. Of course, race dummies were not relevant in the WLS regressions.

results)—but the availability of these covariates varied widely across surveys, and in any case few were statistically significant or materially affected our results.

Tables 7a and 7b provide coefficient estimates from these estimations. Table 7a implements centile rankings as our measure of academic ability, while 7b uses standard scores. The marginal effects of test score on the likelihood of entering teaching—calculated as  $\partial f / \partial x_{it}$  in equation (1)—and their respective elasticities (the marginal effect divided by the mean of  $y_{it}$ ) are also computed for white females with average parental education, both at the mean and one standard deviation above the mean, or at the 85<sup>th</sup> centile.<sup>24</sup>

As one would expect in a sample of graduates with a minimum of a high school degree, we find in Tables 7a-b that test scores and entry into teaching have a positive, statistically significant relationship across all four cohorts (columns 1, 3, 4, and 6). Inclusion of race and parental education controls (columns 2, 5, and 7) in all cases reduces somewhat the estimated coefficient on test scores. Race appears to be important only in the NLS-72 sample, where black women are more likely (all else equal) to become teachers than white women.<sup>25</sup>

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 z-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 z-scores are used as a measure of ability. As seen in Table 7a, the elasticity of a one-centile increase in test score calculated at the mean white female has dropped

---

<sup>24</sup> Elasticities—or, the *percentage* increase in the probability 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 over the four cohorts.

<sup>25</sup> The lack of significance on the race dummy coefficients are most likely due to the comparatively fewer teachers that are members of ethnic minorities, and our use of sampling weights with the NLS-72 and HSB regressions (weights diminish the importance of observations from the oversampled group).

from 0.0192 in 1964 (with a 95 percent confidence interval of [0.0162, 0.0248]), to 0.0126 in 1992 (95 percent confidence interval of [0.0065, 0.0187]), or about a third of its size (compare columns (2) and (7)).<sup>26</sup> These effects are again a bit stronger when using standard scores as our measure of academic ability. Table 7b indicates the elasticity of a one standard deviation increase in test score calculated at the mean white female has dropped almost in half—from 0.5341 (95 percent confidence interval of [0.417, 0.651]) to 0.2745 (95 percent confidence interval of [0.114, 0.435]), or about 0.26 percentage points.<sup>27</sup> Unfortunately, given our samples we are unable to reject the hypothesis that these elasticities differ from 1964 to 1992, at the 95 percent level, but nevertheless find the sharp decline in this relationship between test scores and entry into teaching to be striking.

Results thus far suggest that the positive relationship between academic ability and entry into teaching among female high school graduates has indeed weakened over time. Our stronger results with the use of standard scores also suggest that individuals near the tails of the test score distribution are 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 (1) again for each cohort, but instead implement decile dummies as our measure of ability.

Table 8 presents the results of this estimation (the lowest decile is the omitted group). We again include parental education and race dummies as covariates, where available. Here, most of the coefficients on the decile dummies are positive and statistically significant (in this

---

<sup>26</sup> This result differs little when the elasticity is computed at the 85<sup>th</sup> percentile vs. the mean centile.

<sup>27</sup> Note that a movement of one z-score will be a much greater movement (in magnitude) than one centile rank. Thus the marginal effects and elasticities will correspondingly be larger in magnitude when using z-scores. We are more interested here in the *change* in these effects than the magnitude, per se.

context, statistical significance simply means that being in decile  $j$  makes one more likely to be a teacher than if she were in the lowest decile).

Because marginal effects and elasticities like those provided in Tables 7a-b are less meaningful under this specification, we have calculated the predicted probability that a white female with average parental education in decile  $j$  becomes a teacher, for all surveys (see Table 9). For white females in most deciles the probability of being identified as a teacher (controlling for parental education) has fallen roughly in half from 1964 to 1992 (compare columns (2) and (7) in Table 9). We find however, much larger drops in this probability for females in the top, 9<sup>th</sup>, and bottom deciles, with a greater than fivefold decline in this probability among those in the top decile. Figure 4 plots these predicted probabilities for white females. 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.

If women near the top of their high school classes are increasingly less likely over this period to enter teaching, what career paths were they pursuing? We tabulated the top ten occupations reported by females ranking in the top decile of their class, in all four 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 percent), accountants (6.0 percent), or managers (15.1 percent).<sup>28</sup> Top decile females were almost as likely to be lawyers and judges (3.2 percent) as teachers.

### Male Entrants

An intriguing possible side effect of the gender desegregation of occupations and the movement of talented women into high-cognitive ability occupations is the substitution of high skilled *men* into teaching. As men lost a virtual monopoly on certain professions, finding themselves in competition with capable women for positions, 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 10.

Our results in Table 10 are quite interesting, if only suggestive. Across these four cohorts, we find that the average academic ability of *male* teachers rose from 1964-1992 by 12 percent, or 76 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 three lines in Table 10 indicate. While (as with women) the probability that a male high school graduate entered teaching fell over this period for all decile groups, 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 teaching of 77-84 percent, this reduction was only 57 percent for those in the top decile.

Again, appropriate caution should be used with the results in Table 10. Our sample of male teachers was naturally much smaller than that of females (as few as 60 male teachers in the case of HSB), and the representation of males (particularly top decile males) in the general

---

<sup>28</sup> 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).

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 teachers than elementary teachers, suggesting that if anything, these results are most relevant for male secondary level teachers. Nevertheless, we find our results on men to be intriguing, and worthy of additional study with more suitable data.

## **VII. Conclusion**

Despite the wealth of cross-sectional studies that have examined the characteristics of college graduates choosing teaching as a career, there has been little empirical evidence on how these characteristics—particularly the academic ability of these teachers—have changed over 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 this paper, we have found—at least among these four cohorts of high school graduates—a slight but detectable decline in the academic ability of the average new female teacher (age 25-26), when ability is measured as one's centile rank on a standardized test of verbal and mathematical aptitude. This decline is even greater when measuring ability using standard ( $z$ -) scores. In addition, we find that examination of the entire ability distribution of new teachers is more informative than trends in central tendency alone. Consistent with the earlier findings of Murnane et. al. (1991) and Bacolad (2001) on college graduates in the NLSY, we find that the likelihood of entering teaching as a profession by age 25-26 has fallen for all ability levels of high school graduates, but this probability has fallen remarkably more for females at the top of the distribution—females in the top decile of their high school class are less



than one-fifth as likely to become teachers in 1992 vs. 1964. While our sample sizes of men are much smaller, we detect the opposite trend among men.

While we are confident that these results are indicative of trends in teacher quality among entering female teachers, and believe that our estimates provide some initial measures of the extent to which females of high verbal and math ability are less likely to enter teaching, we acknowledge some shortcomings inherent in our data. First, while our four datasets all have the advantage of being longitudinal in nature (allowing us to compare respondents to others in their high school class *and* be confident of their ultimate occupational choice), they provide snapshots of only four particular moments in time. Second, the exams given to students and the other covariates available varied from one survey to the next. Consequently, we cannot be completely certain that the standardized test scores used here measure exactly the same skills or aptitudes, nor can we consistently control for as many other interesting covariates across the four surveys as we would wish. Finally, despite the importance of the academic ability of teachers to student success, we acknowledge again that test scores are but one measure of teacher “quality,” and do not claim to have captured in this measure the myriad of important skills, talents, and character traits that make for a great teacher.

If our results can be applied to the wider population of new teachers in the United States, a given student in 1992 (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 low income or predominately black

school encounters a teacher of the highest academic ability is likely even lower.<sup>29</sup> For the casual observer, these results will surprise few (as suggested by our lead quote from Chancellor Levy). However, we hope that our quantitative results here will serve as an impetus for further research into changes in teacher quality over time (our curious finding on the quality of male teachers in particular calls out for more work and better data). 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.

---

<sup>29</sup> See, for example, Lankford, Loeb and Wyckoff (2001).

## DATA APPENDIX

### **Wisconsin Longitudinal Study**

The Wisconsin Longitudinal Study (WLS) is perhaps the earliest major longitudinal study of high school graduates. It is a random sample of 10,317 (virtually all non-Hispanic white) men and women who graduated from Wisconsin high schools in 1957. Follow-up surveys collected in 1964, 1975 and 1992 contain, among other things, information about each respondent's schooling, work history, and labor market experiences. Also included are test scores from the Henmon-Nelson test of mental ability. A module containing very thorough information about female work histories was added in 1993. In this paper, we focus primarily on the original 1957 survey and the 1964 follow up, conducted when most respondents were 25 years of age. Contrary to the other follow-up surveys, the 1964 survey was completed by the respondents' parents. Complete documentation for the WLS can be found on-line at <http://dpls.dacc.wisc.edu/WLS/>.

### **Project Talent**

Project Talent is another early longitudinal study of high school students, administered to 400,000 students grades 9-12 in 1960, designed to be a representative sample of American youth. Follow-up surveys were conducted 1, 5, 10, and 20 years following the expected year of graduation from high school. Designed to allow investigation into the relationship between students' cognitive abilities, family and school environment, and post-high school education, work experience, and family development, the study includes measures of numerous aptitudes, and work history. This paper makes use of the Project Talent Public Use File, a 4,000-student sub-sample of the original participants. The sub-sample includes 1,000 students from each graduating class 1960-1963, and was chosen such that every included student was a participant in the 11-year follow up. We focus primarily on this 11-year follow-up, when most respondents were 26-29 years of age. For more information about Project Talent, refer to Flanagan, et. al. (1981).

### **National Longitudinal Study of the High School Class of 1972**

The National Longitudinal Study of the High School Class of 1972 (NLS-72) was the first major nationwide longitudinal study of high school students conducted by the U.S. Department of Education. 16,683 high school seniors were surveyed in 1972, and then contacted again in 1973-74, 1974-75, 1976-77, 1979-80 and 1986. The study includes information on each respondent's family background, community, education, family development, and labor market experience. Also included are several measures of academic ability—SAT and ACT scores (where available), and raw scores on a battery of tests administered by the NLS to most seniors in 1972. The fifth follow up in 1986 included a supplemental survey of all of those in the original 1972 sample who had obtained teaching certificates and/or who had teaching experience. We focus primarily on the fourth NLS-72 follow-up in 1979, when most respondents were 25 years of age. More information on the NLS-72 can be found online at <http://nces.ed.gov/surveys/nls72/>

## High School and Beyond

High School and Beyond (HSB) is another major longitudinal study of high school students conducted by the U.S. Department of Education. It was designed to be comparable to and improve upon the NLS-72. Two cohorts of students—sophomores and seniors—were surveyed and tested in 1980, and then re-surveyed in 1982, 1984, 1986 and 1992. Like the other surveys mentioned above, HSB and its follow-ups include detailed information on each student's family background, school experiences, higher education and training, and career. More detailed academic data (such as GPA, class rank, attendance, courses taken and grades received, etc.) is available for a subset of students in the HSB Transcripts Survey. SAT and ACT scores are available for many students, as are raw scores on the HSB battery of tests, administered to both sophomores and seniors in 1980 (sophomores were re-tested in 1982). We make use only of the *sophomore* cohort survey, and the 1992 follow-up, when most respondents were 28 years of age. More information on HSB can be found at <http://nces.ed.gov/surveys/hsb/>

## Current Population Survey

The Current Population Survey (CPS) is a monthly survey of about 50,000 households, conducted by the Bureau of Labor Statistics and intended to be the primary source of information about the labor force characteristics of the United States. In this paper, we use public use microdata from the CPS March Supplement, or Annual Demographic Survey, 1964-1996. For more information, see <http://www.bls.census.gov/cps/cpsmain.htm>.

## IPUMS

IPUMS is the “integrated public use microdata series,” a number of samples (at various levels of precision) of the U.S. population from the decennial U.S. Census. In this paper, we make use of the 1980 and 1990 5% State IPUMS samples, a 1-in-20 sample of the U.S. population, and the 1970 1% State PUMS, a 1-in-100 sample. See <http://www.ipums.umn.edu/usa/index.html> for more information.

## REFERENCES

- Bacolad, Marigee P. "The Role of Alternative Opportunities in the Female Labor Market on Teacher Supply and Quality 1940-1990." *Mimeo, UCLA Dept. of Economics*. November 2001.
- Ballou, Dale and Michael Podgursky. *Teacher Pay and Teacher Quality*. Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research, 1997.
- Beller, Andrea. "Occupational Segregation by Sex: Determinants and Changes." *Journal of Human Resources*. 17(3): pp. 371-392, 1992.
- Blau, Francine D., Patricia Simpson, and Deborah Anderson. "Continuing Progress? Trends in Occupational Segregation Over the 1970's and 1980's." *Feminist Economics* 4(3): pp.29-71, 1998.
- Ehrenberg, Ronald G. and Dominic J. Brewer. "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond." *Economics of Education Review*. 13(1), pp. 1-17, 1994.
- Ehrenberg, Ronald G. and Dominic J. Brewer. "Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited." *Economics of Education Review*. 14(1), pp. 1-21, 1995.
- Ferguson, Ronald F. "Can Schools Narrow the Black-White Test Score Gap?" in *The Black-White Test Score Gap*, ed. Christopher Jencks and Meredith Phillips. Washington, D.C.: Brookings Institution, 1998.
- Ferguson, Ronald F. and Helen F. Ladd. "How and Why Money Matters: An Analysis of Alabama Schools." *Holding Schools Accountable: Performance-Based Reform in Education*, ed. Helen F. Ladd. Washington, D.C.: Brookings Institution, 1996.
- Flanagan, John C., David V. Tiedeman, William V. Clemans, Laress L. Wise. *Project TALENT Public Use File, 1960-1976*. Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research, 1981.
- Flyer, Frederick and Sherwin Rosen. "The New Economics of Teachers and Education." *Journal of Labor Economics*. 15(1), pp. S104-S139, 1997.
- Gitomer, Drew H. and Andrew S. Latham. "The Academic Quality of Prospective Teachers: The Impact of Admissions and Licensure Data." *Educational Testing Service Research Report*. May 1999.
- Goldhaber, Dan. "The Mystery of Good Teaching: Surveying the Evidence on Student Achievement and Teachers' Characteristics." *Education Next*. 2(1):50-55, 2002.

- Hanushek, Eric A. "The Production of Education, Teacher Quality, and Efficiency." *Do Teachers Make a Difference? A Report on Recent Research on Pupil Achievement*. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1970.
- Hanushek, Eric A. "Teacher Characteristics and Gains in Student Achievement." *American Economic Review*. 61(2), pp. 280-288, 1971.
- Hanushek, Eric A. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*. September 1986, pp. 1141-1177.
- Hanushek, Eric A. "School Resources and Student Performance." *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, ed. Gary B. Burtless. Washington, D.C.: Brookings Institution, 1996.
- Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. "Teachers, Schools, and Academic Achievement." *NBER Working Paper #6691* (revised) 2000.
- Hanushek, Eric A. and Richard R. Pace. "Who Chooses to Teach (and Why)?" *Economics of Education Review* 14(2), pp. 101-117, 1995.
- Krueger, Alan. "Economic Considerations and Class Size." *Princeton University, Industrial Relations Section Working Paper #447*, 2002.
- Lakdawalla, Darius. "The Declining Quality of Teachers." *NBER Working Paper #8263*, 2001.
- Lankford, Hamilton, Susanna Loeb, and James Wyckoff. "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis." *Stanford University Working Paper*, August 2001.
- Lloyd, Cynthia B. and Beth T. Niemi. *The Economics of Sex Differentials*. New York: Columbia University Press, 1979.
- Manski, Charles F. "Academic Ability, Earnings, and the Decision to Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972." *NBER Working Paper #1539*, 1985.
- Murnane, R.J., J.D. Singer, J.B. Willett, J.J. Kemple and R.J. Olsen. *Who Will Teach?* Cambridge: Harvard University Press, 1991.
- Pavalko, Ronald M. "Recruitment to Teaching: Patterns of Selection and Retention." *Sociology of Education*. 43(3), pp. 340-353, 1970.
- Podgursky, Michael, Ryan Monroe and Donald Watson. "Teacher Mobility, Pay, and Academic Quality." *Mimeo, University of Missouri—Columbia*, 2002.
- Polachek, Solomon W. "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics*. 63(1), pp. 60-69, 1981.

Siltanen, Janet, Jennifer Jarman and Robert M. Blackburn. *Gender Inequality in the Labour Market: Occupational Concentration and Segregation*. Geneva: International Labour Office, 1995.

Stoddard, Christiana. "Why Has the Number of Teachers Per Student Risen While Teacher Quality Has Declined?" *Mimeo, Department of Economics, UC Santa Barbara*. December 2001.

Temin, Peter. "Teacher Quality and the Future of America." *NBER Working Paper #8898*, 2002.

Vance, Victor S. and Philip C. Schlechty. "The Distribution of Academic Ability in the Teaching Force: Policy Implications." *Phi Delta Kappan*. September, 1982, pp. 22-27.

Vegas, Emiliana, Richard J. Murnane, and John B. Willett. "From High School to Teaching: Many Steps, Who Makes It?" *Teachers College Record*. 103(3), June 2001, pp. 427-449.

Wayne, Andrew J. and Peter Youngs. "Teacher Characteristics and Student Achievement Gains: A Review." *Mimeo, SRI International*, 2001.

Weaver, W. Timothy. *America's Teacher Quality Problem: Alternatives for Reform*. New York: Praeger Publishers, 1983.

**Figure 1: Females 25-34--Percent Teachers**

Source: March CPS 1964-1996

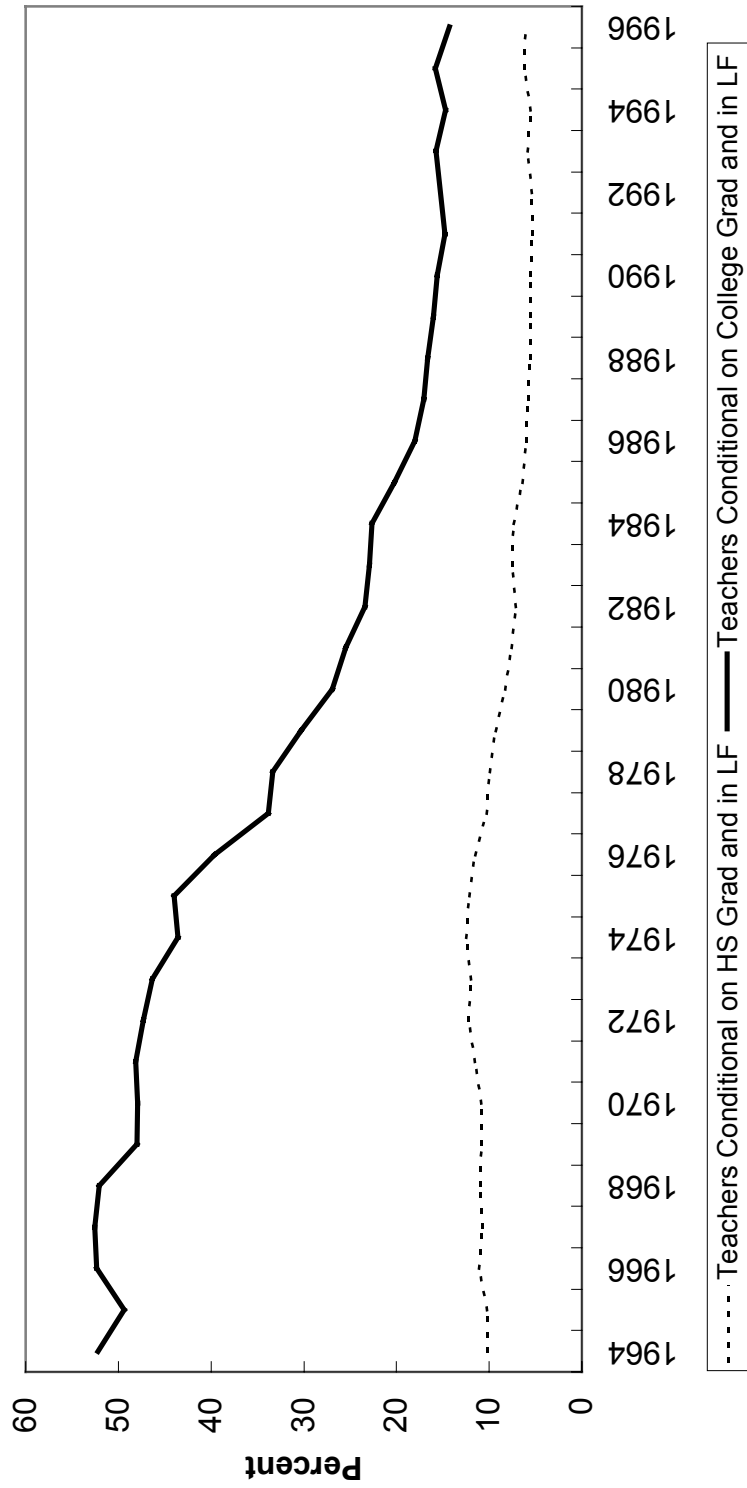




Figure 2: Female College Completion Rates by Test Score Decile

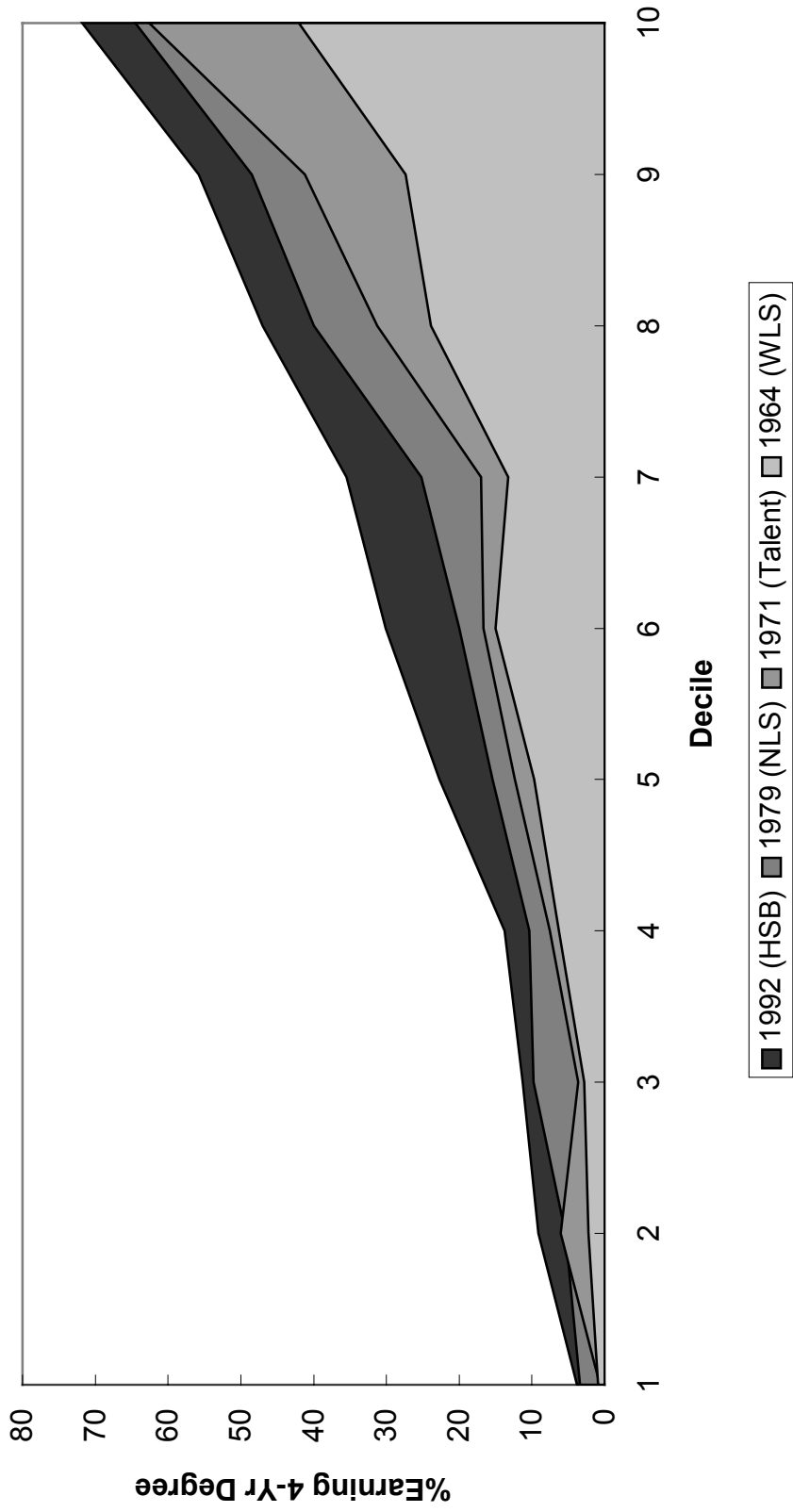
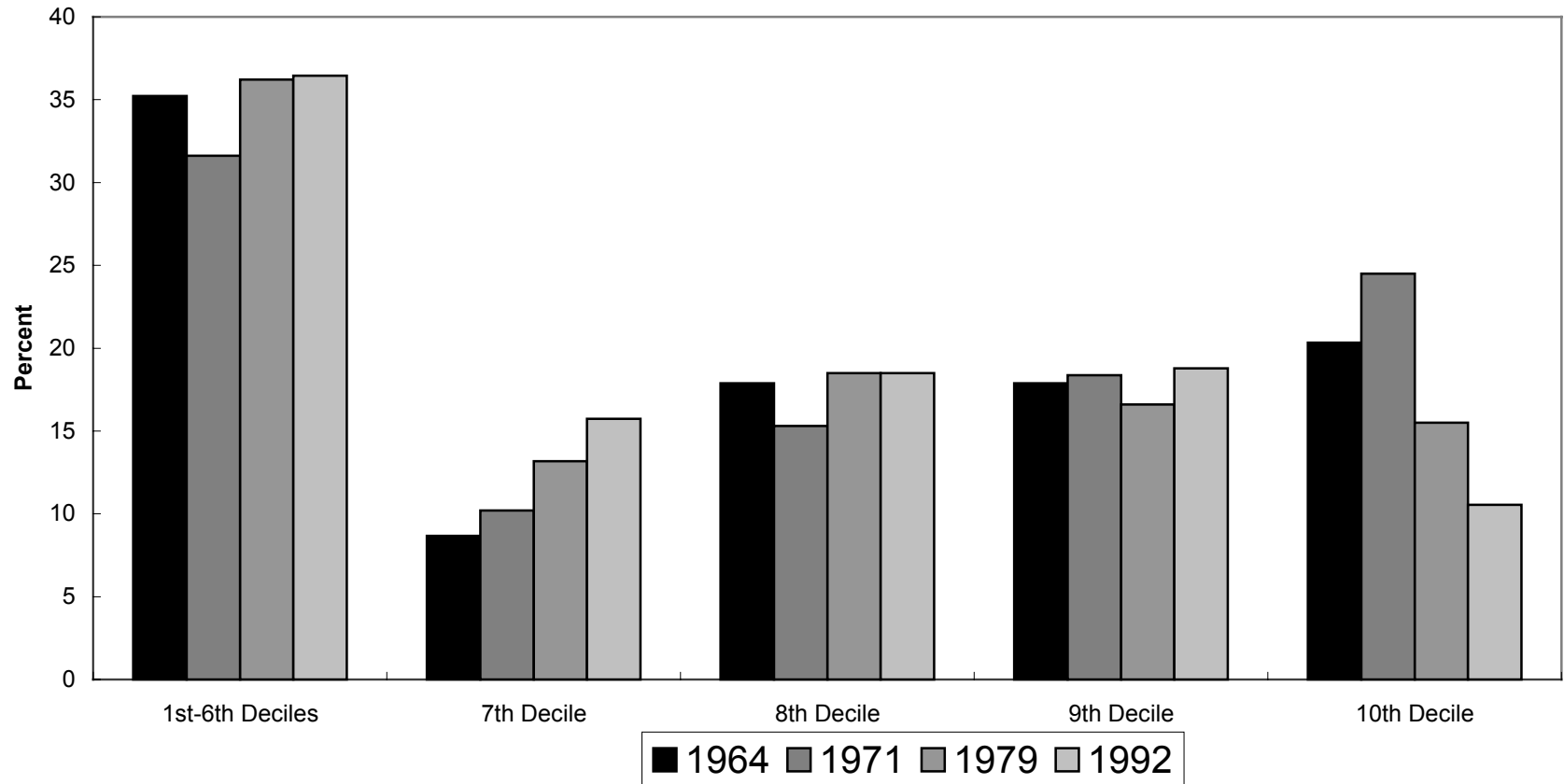


Figure 3: Distribution of Teachers Across Test Score Decile Groups 1964-1992



**Figure 4: Predicted Probability of Becoming a Teacher, by Decile  
(White Female HS Graduates)**

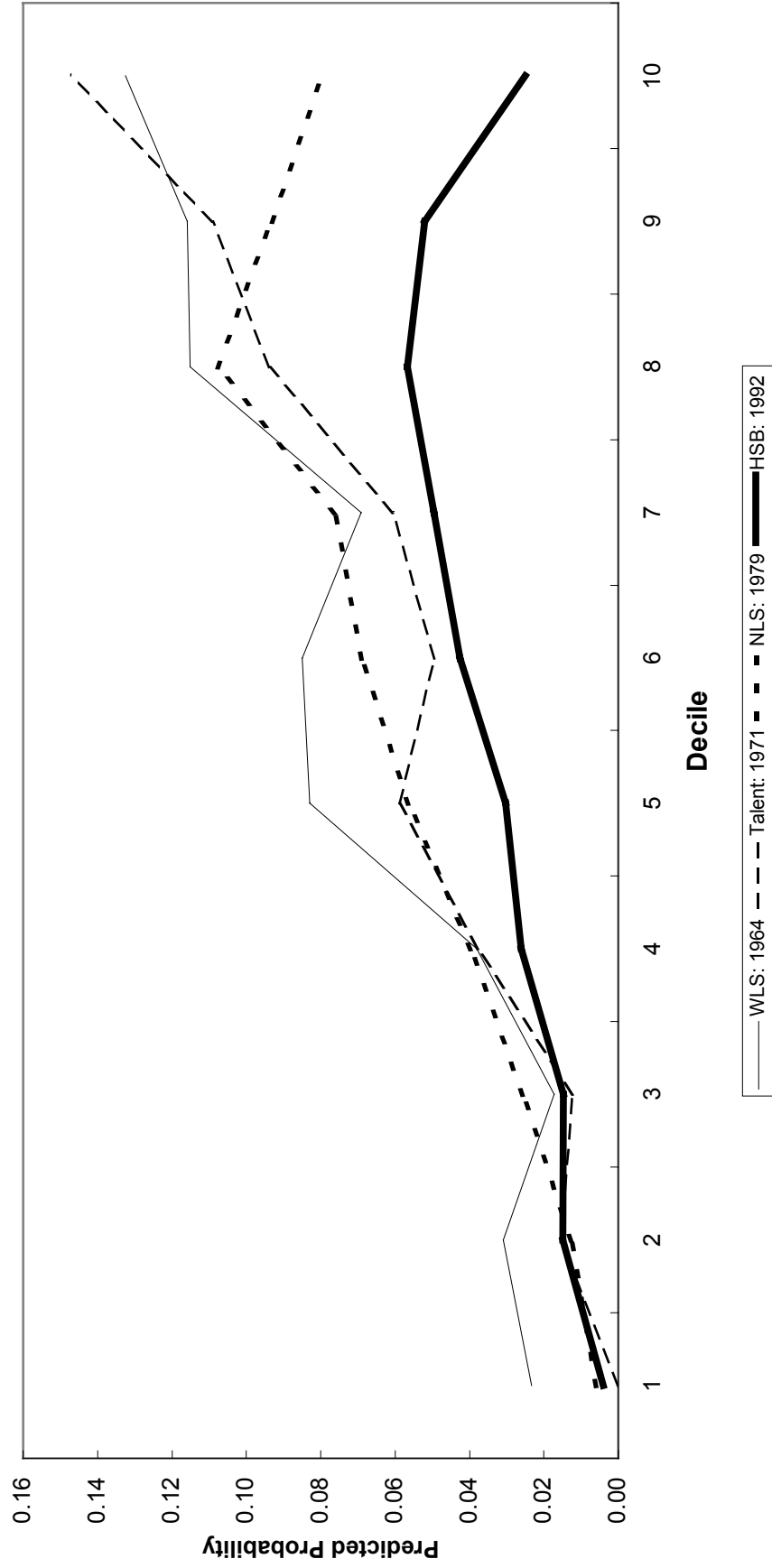


Table 1: Labor Market Characteristics, Selected Years, 1964-1996

	1964	1970	1980	1990	1996
% Young Females (age 25-34) with:					
At least a high school degree	67.4	73.3	84.3	86.5	86.5
At least four-year college degree	8.9	12.0	20.7	22.8	26.1
In the labor force	37.2	45.7	65.4	73.2	74.0
Index of gender representation (age 25-34)					
Physicians	0.33	0.20	0.47	0.69	0.79
Lawyers	0.08	0.20	0.44	0.60	0.94
Public elementary teachers (millions)	1.0	1.3	1.4	1.7	1.9
Public secondary teachers (millions)	0.6	1.0	1.1	1.1	1.2
Public school enrollment (millions)	39.0	52.2	49.7	53.8	59.5
% Young teachers (age 25-34) that are female					
All	--	70.0	72.4	74.2	76.3
Primary	--	85.1	84.6	86.0	82.8
Secondary	--	46.6	54.4	53.0	57.2

Source: March Current Population Survey 1964-1996, except teacher counts, taken from the Statistical Abstract of the United States, 2000. First teacher count figure for 1960, first enrollment is from 1965. Gender representation index is calculated from data beginning in 1968.

Table 2: Descriptive Statistics for Female High School Graduates

	WLS	Talent	NLS-72	HSB
High school graduation year(s)	1957	1960-63	1972	1982
Follow-up survey year(s)	1964	1971-74	1979	1992
Sample size <sup>1</sup>	4,609	1,633	6,752	5,360
Response rate of follow-up survey (%)	86.5	100.0	88.5	90.7
Average age in years at follow-up	25.0	26.8	26.1	27.5
Race/Ethnicity (%):				
White	100.0	88.1	81.2	79.6
Black	--	6.4	10.1	12.7
Hispanic	--	1.3	3.5	5.6
Other	--	2.6	5.2	2.2
% With at least 2 year of college	28.6	29.6	32.0	36.4
% With at least 4 years of college	14.5	19.9	24.2	27.2
% Working full or part-time	38.4	47.1	72.0	77.1
Mean, mother's education (years)	10.3	--	12.4	12.7
Mean, father's education (years)	9.7	--	12.7	13.0
Mean, centile rank <sup>2,3</sup>				
High school graduates	50.1	49.5	48.9	49.7
College graduates	73.4	74.2	71.1	71.8
Mean, standard score				
High school graduates	0.019	0.002	-0.020	-0.008
College graduates	0.820	0.855	0.729	0.793

<sup>1</sup> 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. Sampling weights were used in generating means in NLS-72 and HSB surveys.

<sup>2</sup> Centile ranks are based on students' placement in the distribution of all graduates of the same gender in their high school class. For the NLS-72 and HSB surveys, centile ranks were assigned using an algorithm incorporating base-year weights.

<sup>3</sup> 'College graduates' refers to completion of a four-year bachelor's degree.

Table 3: Comparison of White Female High School Graduates Born in Wisconsin versus Other Women, 1936-1942 Cohorts, 1970, 1980 and 1990 PUMS

	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 <sup>1</sup>	WI	All other areas <sup>1</sup>	WI	All other areas <sup>1</sup>		
Sample size	1,494	51,174	7,475	271,924	7,939	295,978		
% Unemployed	3.85	3.75	3.55	3.73	2.74	3.14		
% Not in labor force	56.5	56.4	29.2	32.3	22.8	26.8	**	**
Teachers (% among high school grads)	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	**	**
Bachelor's degree (%) <sup>2</sup>	16.5	16.7	16.8	18.5	17.9	19.8	**	**
Master's degree (%) <sup>2</sup>	1.4	2.2	3.9	4.7	5.95	8.0	**	**
Real wage/salary income (1990 \$) <sup>3</sup>	15,262 (8,509)	16,463 (8,647)	15,345 (9,919)	16,677 (10,367)	18,455 (13,884)	20,897 (15,008)	**	**
Real weekly wage (1990 \$) <sup>3</sup>	310.19 (174.90)	333.80 (175.52)	307.54 (198.49)	334.64 (207.48)	369.35 (277.02)	418.24 (300.42)	**	**
% Living in MSA	59.7	64.9	69.2	69.4	56.4	63.5	**	**
% Living in same state as birth	67.5	57.9	65.6	56.4	65.4	55.6	**	**

\*\* Identifies a statistically significant difference in sample means. Person weights were used in the 1990 samples. The 1970 and 1980 samples were designed to be self-weighting.

<sup>1</sup> Females from all other areas may include foreign born.

<sup>2</sup> In the 1970 and 1980 samples, we assume those with at least four years of college have a bachelor's degree, and those with at least six years of college have a master's degree.

<sup>3</sup> In computing mean wages or income, only those earning at least \$1000/year and working at least 40 weeks/year were included.

Table 4: Logit Estimates of the Probability of Male High School Graduates Entering Medicine

Sample	WLS		NLS-72		HSB	
Dependent variable	Physician in 1975		Physician in 1986		Physician in 1992	
Mean of dependent variable	0.011		0.023		0.008	
Mean age of respondent	43.0		32.9		27.4	
Sample size	4,331		4,358		4,834	
Variable	(1)	(2)	(1)	(2)	(1)	(2)
Constant	-8.075 (0.696)	-5.347 (1.374)	-7.323 (0.502)	-4.760 (0.206)	-8.567 (0.790)	-5.671 (0.286)
Centile rank	0.052 (0.008)		0.051 (0.006)		0.055 (0.010)	
Standardized score		1.374 (0.198)		1.473 (0.167)		1.330 (0.198)
Marginal effect of test score at:						
Mean	0.0002	0.0065	0.0004	0.0124	0.0002	0.0045
95 <sup>th</sup> percentile or +1 SD	0.0013	0.0249	0.0023	0.0512	0.0011	0.0169
Elasticity of marginal effect at:						
Mean	0.0521	1.3676	0.0506	1.4603	0.0546	1.3250
95 <sup>th</sup> percentile or +1 SD	0.0509	1.3487	0.0486	1.4197	0.0537	1.3125

Standard errors in parentheses. For the WLS, NLS-72 and HSB surveys, we used data from the 1975, 1986 and 1992 follow-up respectively.

Table 5: Most Frequently Reported Occupations

A: Females with a High School Degree

Rank	WLS -- 1964		Talent -- 1971-74		NLS-72 -- 1979		HSB -- 1992	
	Occupation	%	Occupation	%	Occupation	%	Occupation	%
1	None/housewife	27.0	Housewife	45.1	Clerical	35.4	Clerical	34.2
2	Clerical	15.9	Secretary	4.4	Service worker	11.9	Service worker	16.3
3	Stenographer/typist	12.8	Elementary teacher	2.7	Manager	8.7	Manager	11.0
4	Teacher	9.7	Registered nurse	2.2	Operative (non transport)	7.1	Sales/worker	5.3
5	Nurse/professional	4.6	Student	1.8	Teacher	6.7	None/housewife	3.7
6	Bookkeeper	2.9	Bookkeeper	1.6	Sales	6.1	Operative (non transport)	3.6
7	Office machine operator	2.1	Factory assembly	1.4	Nurse, dietician, therapist	5.0	Teacher	3.5
8	Waitress	1.9	Clerk/typist	1.2	None/housewife	3.0	Craftsman	2.6
9	Beautician	1.9	None	1.1	Health technician	2.1	Health technician	2.2
10	Medical/dental technician	1.8	Clerical -- other	1.0	Craftsman	2.0	Nurse, dietician, therapist	2.2

B: Females with at least a Four-Year Degree

Rank	WLS -- 1964		Talent -- 1971-74		NLS-72 -- 1979		HSB -- 1992	
	Occupation	%	Occupation	%	Occupation	%	Occupation	%
1	Teacher	48.8	Housewife	34.1	Teacher	24.6	Clerical	17.3
2	None/housewife	14.2	Elementary teacher	13.1	Clerical	16.4	Manager	13.9
3	Student	3.9	Teacher - other	3.5	Manager	11.5	Teacher	11.8
4	Social worker	3.4	Student	2.6	Nurse, dietician, therapist	9.2	Sales	8.9
5	Stenographer/typist	3.1	Registered nurse	2.3	Sales	4.9	Service worker	6.7
6	Nurse, professional	3.1	Secondary teacher	2.3	Service worker	4.3	Accountant	5.4
7	Medical/dental technician	2.5	Social worker	2.0	Writer/artist	3.2	Nurse, dietician, therapist	5.1
8	College professor	2.2	Sec. teacher/P.E.	1.5	Health technician	2.6	Computer specialist	4.2
9	Therapist/healer	2.1	Music teacher	1.5	Social worker	2.1	Writer/artist	4.1
10	Musician/music teacher	1.5	Pre-K/K teacher	1.2	Accountant	2.0	Health technician	2.3

Project Talent differed from other surveys in that women who were out of the labor force did not report an occupation. Consequently, many of the women who reported “housewife” in Project Talent would have specified a profession in any of the surveys. Also, Project Talent did not use Census classifications of occupations, as the other surveys did, reducing comparability between occupation titles in this table.



Table 6: Descriptive Statistics of New Female Teachers

	WLS	Talent	NLS-72	HSB
High school graduation year	1957	1960-63	1972	1982
Follow-up survey year	1964	1971-74	1979	1992
Number of teachers <sup>1</sup>	369	98	430	219
% High school graduates	8.0	6.0	6.7	4.1
% College graduates	55.2	30.2	25.5	13.3
Race of teacher (%)				
White	100.0	90.8	85.4	88.0
Black	--	6.1	9.2	8.2
Hispanic	--	--	1.5	3.4
Other	--	3.1	3.9	0.5
Ability measures:				
Mean centile rank <sup>2</sup>	67.3	69.3	65.1	64.4
Mean standard score <sup>3</sup>	0.60	0.69	0.53	0.48

<sup>1</sup>Teachers were identified using the following codes in each survey. WLS (#51 – Teachers, NEC”), Project Talent (#400-433 – all elementary and secondary teachers), NLS-72 (#142-145 – “Elementary Kindergarten, Secondary, and Teachers except college and university, nec). Occupations reported by the sophomore cohort in HSB were not codified using any numeric system, so we hand-coded each respondent’s occupation.

<sup>2</sup>For the NLS-72 and HSB surveys, centile ranks were assigned using an algorithm incorporating base-year weights. Mean centile ranks calculated using follow-up weights.

<sup>3</sup>Scores were standardized using base-year survey weights; mean standardized scores calculated using follow-up weights.

Table 7a: Logistic Regressions of Teacher Entrance Models for all Female High School Graduates

Dependent variable	WLS Teacher in 1964		Talent Teacher in 1971	NLS-72 Teacher in 1979		HSB Teacher in 1992	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample mean	0.080		0.060	0.064		0.041	
Sample size	4,609		1,633	6,752		5,360	
Intercept	-3.885 (0.153)	-5.192 (0.254)	-4.427 (0.310)	-3.965 (0.135)	-5.803 (0.358)	-4.412 (0.194)	-6.634 (0.460)
Centile score	0.025 (0.002)	0.021 (0.002)	0.022 (0.002)	0.022 (0.002)	0.021 (0.002)	0.019 (0.003)	0.013 (0.003)
Mother's years of education		0.071 (0.023)			0.121 (0.029)		0.096 (0.038)
Father's years of education		0.076 (0.018)			0.024 (0.025)		0.099 (0.032)
Black					0.625 (0.198)		0.026 (0.287)
Hispanic					-0.100 (0.427)		-0.156 (0.415)
Other race					-0.090 (0.270)		-1.515 (1.051)
Marginal effect of test score at:							
Mean centile	0.0015	0.0012	0.0013	0.0014	0.0011	0.0007	0.0004
85 <sup>th</sup> percentile	0.0031	0.0022	0.0029	0.0026	0.0018	0.0012	0.0006
Elasticity of marginal effect at:							
Mean centile	0.0231	0.0192	0.0268	0.0232	0.0195	0.0201	0.0126
85 <sup>th</sup> percentile	0.0212	0.0180	0.0250	0.0242	0.0186	0.0205	0.0124

Standard errors in parentheses. Sample weights used with the NLS-72 and HSB samples. Marginal effects and elasticities calculate for a white female with an average test score and parental education.

Table 7b: Logistic Regressions of Teacher Entrance Models for all Female High School Graduates

Dependent variable	WLS Teacher in 1964		Talent Teacher in 1971	NLS-72 Teacher in 1979		HSB Teacher in 1992	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample mean	0.080		0.060	0.064		0.041	
Sample size	4,609		1,633	6,752		5,360	
Intercept	-2.646 (0.065)	-4.185 (0.241)	-3.025 (0.131)	-2.888 (0.059)	-4.775 (0.360)	-3.416 (0.082)	-6.070 (0.475)
Standardized score	0.698 (0.062)	0.570 (0.064)	0.816 (0.119)	0.662 (0.059)	0.624 (0.065)	0.483 (0.074)	0.284 (0.085)
Mother's years of education		0.072 (0.023)			0.120 (0.025)		0.099 (0.032)
Father's years of education		0.077 (0.018)			0.024 (0.029)		0.105 (0.038)
Black					0.651 (0.198)		-0.066 (0.284)
Hispanic					-0.085 (0.427)		-0.213 (0.414)
Other race					-0.082 (0.271)		-1.547 (1.050)
Marginal effect of test score at:							
Mean centile	0.0437	0.0340	0.0391	0.0411	0.0322	0.0173	0.0091
85 <sup>th</sup> percentile	0.0762	0.0542	0.0727	0.0707	0.0501	0.0256	0.0113
Elasticity of marginal effect at:							
Mean centile	0.6457	0.5341	0.7748	0.7007	0.5895	0.4992	0.2745
85 <sup>th</sup> percentile	0.6106	0.5098	0.7353	0.7251	0.5685	0.5067	0.2721

Standard errors in parentheses. Sample weights used with the NLS-72 and HSB samples. Marginal effects and elasticities calculate for a white female with an average test score and parental education.

Table 8: Logistic Regressions, Who Enters Teaching, Females with a High School Degree

Dependent variable	WLS		Talent	NLS-72		HSB	
	Teacher in 1964		Teacher in 1971	Teacher in 1979		Teacher in 1992	
Sample mean	0.080		0.060	0.064		0.041	
Sample size	4,609		1,633	6,742		5,360	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	-3.857 (0.450)	-5.237 (0.396)	-4.159 (0.451)	-4.887 (0.448)	-7.130 (0.571)	-5.610 (0.733)	-8.227 (0.859)
10 <sup>th</sup> decile	2.261 (0.493)	1.858 (0.365)	2.402 (0.502)	2.626 (0.466)	2.690 (0.483)	2.348 (0.767)	1.828 (0.785)
9 <sup>th</sup> decile	2.014 (0.501)	1.705 (0.366)	2.059 (0.515)	2.731 (0.465)	2.863 (0.480)	2.961 (0.753)	2.589 (0.768)
8 <sup>th</sup> decile	1.985 (0.521)	1.697 (0.365)	1.891 (0.526)	2.869 (0.464)	3.026 (0.477)	2.949 (0.753)	2.680 (0.766)
7 <sup>th</sup> decile	1.334 (0.502)	1.361 (0.374)	1.418 (0.556)	2.476 (0.469)	2.643 (0.481)	2.794 (0.756)	2.535 (0.769)
6 <sup>th</sup> decile	1.548 (0.511)	1.138 (0.386)	1.201 (0.579)	2.338 (0.472)	2.537 (0.483)	2.510 (0.763)	2.377 (0.773)
5 <sup>th</sup> decile	1.441 (0.544)	1.335 (0.377)	1.386 (0.556)	2.111 (0.476)	2.320 (0.487)	2.148 (0.774)	2.022 (0.782)
4 <sup>th</sup> decile	0.634 (0.517)	0.507 (0.419)	0.920 (0.613)	1.810 (0.485)	1.950 (0.491)	1.942 (0.783)	1.872 (0.790)
3 <sup>rd</sup> decile	-0.192 (0.517)	-0.306 (0.492)	-0.235 (0.842)	1.357 (0.502)	1.497 (0.507)	1.306 (0.824)	1.290 (0.828)
2 <sup>nd</sup> decile	0.351 (0.674)	0.291 (0.440)		0.657 (0.553)	0.780 (0.555)	1.310 (0.822)	1.298 (0.823)
Father's years of education		0.078 (0.023)			0.0283 (0.025)		0.107 (0.032)
Mother's years of education		0.072 (0.018)			0.129 (0.029)		0.104 (0.038)
Black					0.849 (0.201)		0.222 (0.288)
Hispanic					0.003 (0.430)		-0.077 (0.415)
Other race					0.0290 (0.271)		-1.449 (1.052)

Standard errors in parentheses. Sample weights used with the NLS-72 and HSB samples. Individuals in the first and second deciles make up the omitted group in the Project Talent models.

Table 9: Predicted Probabilities of Entering Teaching as an Occupation,  
White Females with a High School Degree

Decile of Test score	WLS		Talent	NLS-72		HSB	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
10 <sup>th</sup>	0.169	0.132	0.147	0.096	0.080	0.037	0.025
9 <sup>th</sup>	0.137	0.116	0.109	0.104	0.093	0.066	0.052
8 <sup>th</sup>	0.133	0.115	0.094	0.117	0.108	0.065	0.057
7 <sup>th</sup>	0.074	0.069	0.061	0.082	0.076	0.056	0.050
6 <sup>th</sup>	0.090	0.085	0.049	0.072	0.069	0.043	0.043
5 <sup>th</sup>	0.082	0.083	0.059	0.059	0.056	0.030	0.030
4 <sup>th</sup>	0.038	0.038	0.038	0.044	0.040	0.025	0.026
3 <sup>rd</sup>	0.017	0.017	0.012	0.028	0.026	0.013	0.015
2 <sup>nd</sup>	0.029	0.031	0.015	0.014	0.013	0.013	0.015
1 <sup>st</sup>	0.021	0.023		0.007	0.006	0.004	0.004

Predicted probabilities evaluated for a white female with average parental education in a given decile. Column numbers correspond to reported results in Table 8.

Table 10: Characteristics of New Male Teachers, 1964-1992

	WLS	Talent	NLS-72	HSB	% Change
Sample size (among high school graduates)	4,379	1,527	6,451	4,834	--
Number of teachers	154	68	132	60	--
Mean centile ranking	60.3	59.5	59.9	67.3	11.6
Mean standard score	0.39	0.32	0.37	0.68	75.9
Average predicted probability of teaching					
1 <sup>st</sup> – 5 <sup>th</sup> deciles	0.025	0.039	0.029	0.010	-84.0
6 <sup>th</sup> – 9 <sup>th</sup> deciles	0.043	0.057	0.029	0.010	-76.7
10 <sup>th</sup> decile	0.055	0.052	0.025	0.024	-57.1

Predicted probabilities are estimated in the same manner as those for women in Table 8. Probabilities have been computed for a white male with average parental education.