Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957–2000

By SEAN P. CORCORAN, WILLIAM N. EVANS, AND ROBERT M. SCHWAB*

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 (2000)

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 intensifies. A popular explanation for these frustrations, outlined in Peter Temin (2002), points to the remarkable gender desegregation of the labor market since 1960. 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.

Frustration over the quality of the teaching force comes amidst a growing body of evidence that finds that certain measures of teacher "quality," in particular, their verbal and mathematical skills, are strongly related to student outcomes (see e.g., Ronald G. Ehrenberg and Dominic J. Brewer, 1995; Ronald F. Ferguson and Helen F. Ladd, 1996).[1] The robustness of this finding stands in sharp contrast to the continuing debate over the importance of other inputs into the production of education (notably, per-pupil expenditure and class size).

While the hypothesis that 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. A number of cross-sectional studies (e.g., Charles F. Manski, 1985; Erik A. Hanushek and Richard R. Pace, 1995) have shown that college graduates entering teaching in the 1970's and 1980's did not compare favorably to their peers, but less is known about how this relationship has changed over time. This is largely due to a lack of data. To our knowledge, there are no single data sets currently available that (i) collect data before and after the transformation of the labor market that began in the 1960's, (ii) contain variables that can be considered reliable measures of teacher quality, and (iii) identify practicing teachers with a high degree of certainty.[2] One possible exception is the set of National Longitudinal Surveys, originated in 1968, which Richard J. Murnane et al. (1991) and Marigee P. Bacolod (2003) use to study the changing characteristics of teachers over time. Both find a decline in the fraction of college graduates with

[†] *Discussants:* Brian Jacob, Harvard University; Thomas Kane, University of California–Los Angeles; Julian Betts, University of California–San Diego.

* Corcoran: Department of Economics, California State University, 6000 J Street, Sacramento, CA 95819 (e-mail: corcoran@csus.edu); Evans and Schwab: Department of Economics, University of Maryland, College Park, MD 20742 (e-mail: evans@econ.umd.edu; schwab@econ.umd.edu). We thank the NSF for its financial support. The full version of this paper is available from the authors upon request.

[1] Andrew J. Wayne and Peter Youngs (2003) and Jennifer King Rice (2003) review recent literature on the relationship between teacher characteristics and student achievement.

[2] This absence was recognized in a recent paper by Michael Podgursky et al. (2002 p. 4), 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 ... . Unfortunately, time-series data on teacher quality are not available to directly test this hypothesis."

high AFQT (armed-forces qualifying test) scores choosing to teach.

In this paper, we address this issue by combining five longitudinal surveys of high-school students spanning more than four decades and examine how the propensity for women with high verbal and mathematical skills to become teachers has changed over time. The two data sets that date prior to 1965 afford 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.

I. Data and Methodology

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 for the classes of 1960–1964, the National Longitudinal Study of the High School Class of 1972 (NLS-72), High School and Beyond (HSB) for the class of 1982, and the National Education Longitudinal Study of 1988 (NELS) for the class of 1992. These five surveys are alike in that they each include results from a questionnaire administered during the senior year, all require participation in a battery of aptitude tests, and all conduct numerous follow-ups after graduation. The inclusion of standardized test scores for all students allows us to place graduates into a cohort skill distribution and to assess how the propensity for women (or men) with high relative scores to enter teaching has changed over time. We accomplish this by estimating a series of logit models, wherein we estimate the likelihood that a graduate of each cohort enters the teaching profession, conditional on his or her test score ranking.[3]

The use of large longitudinal surveys of high-school graduates has a number of advantages. First, follow-up surveys allow us to identify those individuals who actually become teachers. Some past attempts to compare teachers with non-teachers (e.g., Dale Ballou and Podgursky, 1997) look at the SAT or ACT scores of students intending to major in education—likely a noisy measure of the skills of practicing teachers, since not all SAT/ACT-takers attend college, and those who do may switch majors, never complete college, or enter the teaching force. Second, they provide us with large samples from a relatively stable population over this period: the population of high school graduates. Comparing teachers to non-teachers in the population of college graduates over this period (as in Murnane, et al. [1991] and Bacolod [2003]) may be misleading if the skill composition of female college graduates has changed over time.

This collection of data is not, however, without its imperfections. For one, these surveys provide snapshots of only five cohorts over this period. In addition, our oldest data set (WLS) consists exclusively of non-Hispanic whites in Wisconsin—clearly not a nationally representative sample of high-school graduates. In the full-length version of this paper, we compare the 1936–1942 birth cohort of Wisconsin-born females with white females born outside of Wisconsin in the 1970 Census Public Use Microdata Sample, and find little evidence to suggest that white women born in Wisconsin look markedly different from white women born elsewhere. The inclusion of Project Talent (which is nationally representative and also pre-1965) should provide a meaningful comparison. Finally, as with all longitudinal surveys, our data are subject to sample attrition, although participation in the follow-up surveys was high.

As mentioned above, our skill measure is a centile ranking based on a student's placement in the distribution of high-school graduates on the math and verbal portions of an exam administered during their senior year. While the specific tests administered by 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 SAT/ACT score, the correlation between the scores is quite high (0.84–0.86).[4] Of course, this single measure of

[3] These models represent the reduced-form relationship between ability and entry into teaching. We do not explore here the channels through which this relationship may have been shaped. In a recent paper, Bacolod (2003) relates changes in relative wages to changes in the quantity and quality of teachers.

[4] As another rough test of the comparability of test scores across surveys, we looked at the relationship between

ability is far from a perfect measure of "teacher quality." Ideally, we would like a vector of those individual characteristics that have positive, causal effects on outputs of the educational process. This vector would likely include many attributes (e.g., patience, creativity, or communication skills) that are unobservable or are otherwise difficult to measure. While we acknowledge the shortcomings of our measure, the empirical literature mentioned in the Introduction appears to support the view that, among observable characteristics, a teacher's math and verbal skills are among the most important for student performance.[5]

We select from each cohort all women who graduated from high school, have a test score available, and responded to a selected follow-up survey. Teachers and non-teachers are identified in these data sets using 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 occupations like teaching (as was suggested by Frederick A. Flyer and Sherwin Rosen [1997]). In this case, a random sample of *working* females would be more likely to exclude teachers than non-teachers. To the extent that labor-force participation among teachers is correlated with academic abilities, this approach would likely result in selection bias.

TABLE 1—DESCRIPTIVE STATISTICS OF FEMALE GRADUATES AND TEACHERS

Statistic	WLS	Talent	NLS72	HSB	NELS
HS graduation year(s)	1957	1960–1963	1972	1982	1992
Follow-up year(s)	1964	1971–1974	1979	1992	2000
Sample size[a]	4,609	1,634	6,751	5,389	4,284
Percentage, 4+ years of college	14.5	20.5	24.7	26.0	42.3
Percentage white	—	88.1	82.1	74.5	74.6
Number of teachers	369	99	431	219	302
Percentage of all female HS graduates	8.0	6.1	6.8	4.1	7.0
Percentage of all female college graduates	55.2	30.4	24.4	13.3	16.3
Percentage white	—	89.9	86.5	80.2	83.8
Mean, centile rank[b]					
HS graduates	50.0	50.4	50.4	50.7	50.5
College graduates	73.4	74.9	72.4	76.2	67.4
Teachers	67.2	69.5	66.4	64.8	63.7
Mean, standard score					
HS graduates	0.02	0.00	0.00	0.01	0.00
College graduates	0.82	0.84	0.74	0.91	0.58
Teachers	0.60	0.65	0.55	0.50	0.46

[a] The sample consists of females who have completed high school by the given follow-up year, responded to the follow up, and had a valid test score. Sampling weights were used in generating means in NLS-72, HSB, and NELS surveys.

[b] Centile ranks are based on students' placement in the distribution of all female graduates in their high school cohort. For the NLS-72, HSB, and NELS surveys, centile ranks were assigned using an algorithm incorporating base-year weights.

II. Results

Table 1 presents some descriptive statistics for female teachers, high-school graduates, and college graduates in each cohort. Not surprisingly, the centile ranking of the average female teacher lies consistently above that of the average high-school graduate (50 by definition, with some variation due to sample attrition). We find, however, that the average female teacher in our sample scored consistently *below* the average female college graduate throughout this period. In addition, the rank of the average female teacher in our sample fell about 3 points over this period, from the 67th centile in WLS (or 69th in Project Talent) to the 64th in the NELS—a drop of 5.2 percent. As centiles mask information about the tails of the distribution, we also computed the mean standard score 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 (0.65 in Project Talent) to 0.46 in 2000, a drop of 23 percent.

centiles and entry into medicine among *men* in these cohorts. Assuming that the relationship between cognitive skills and entry into medicine has changed little since 1960, we would expect to see a consistent relationship between these variables over time, if our test scores measure similar aptitudes. This was indeed the case. See Corcoran et al. (2004) for details.

[5] Darius N. Lakdawalla (2001) and Christiana Stoddard (2003) interpret changes in the relative wage of teachers as evidence of a decline in teacher quality. The link between teacher pay and teacher quality, however, is a tenuous one (see Ballou and Podgursky, 1997; Hanushek et al., 1999). Susanna Loeb and Marianne E. Page (2000) provide an alternative view.

While this modest drop in the mean relative skill of new female teachers 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 decile, age, and race.[6]

In panel A of Table 2, we calculate the (average) predicted probability that a white female in each decile becomes a teacher, for each of the five cohorts. For females in most deciles the probability of being identified as a teacher fell roughly by half from 1964 to 1992, with a modest rise between 1992 and 2000 (compare columns 1, 4, and 5). We find however, much larger drops in this probability for females in the top three deciles from 1964–1992, or the top, 8th, and bottom deciles from 1964–2000 (the NELS cohort alters the pattern somewhat). Panel B of Table 2 divides these predicted probabilities by the overall sample mean in each cohort. Here the trends are clearer: women in the top decile are much less likely to become teachers, relative to the average, in later cohorts versus earlier cohorts. The opposite trend is true for deciles near the bottom of the distribution (with the lowest decile being a notable exception).

Figure 1 illustrates the impact of these differential trends in entry into teaching on the skill composition of female teachers in our samples. In the 1964–1971 period, 20–25 percent of all new female teachers ranked in the top (10th) decile of their high-school cohort; by 2000, this proportion dropped below 13 percent. An increasing share of female teachers scored in the second through sixth deciles in math and verbal abilities; the fraction of female teachers scoring in the lowest (first) decile dropped by half.

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. While our sample sizes are much smaller for male teachers, we repeated our analysis for the five cohorts of male graduates in our surveys. Our results are quite interesting, if only suggestive. Across these cohorts, we find that the test-score ranking of the average male teacher *rose* during 1964–2000 by 6.6 percent. This increase also appears to be driven by those at the top of the distribution: while (as with women) the probability that any male graduate entered teaching fell over this period, 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 35–75 percent from 1964 to 2000, this reduction was only 29

[6] In the full version of this paper, we also allow centiles to enter linearly into the model. Such a specification, which estimates how the strength of the relationship between skill and entry into teaching has changed over time, allows for comparison with Murnane et al. (1991) and others.

TABLE 2—PREDICTED PROBABILITIES OF ENTERING TEACHING, FEMALES WITH AT LEAST A HIGH-SCHOOL DEGREE

A

Decile of test score	Predicted probabilities				
	WLS	Talent	NLS-72	HSB	NELS
10	0.169	0.147	0.096	0.057	0.079
9	0.135	0.111	0.109	0.054	0.145
8	0.122	0.092	0.117	0.046	0.069
7	0.094	0.061	0.089	0.062	0.112
6	0.090	0.049	0.079	0.041	0.089
5	0.079	0.063	0.068	0.047	0.062
4	0.045	0.037	0.048	0.024	0.071
3	0.021	0.012	0.029	0.021	0.049
2	0.024	0.018	0.018	0.022	0.019
1	0.022	0.019	0.001	0.017	0.007
Sample mean:	0.080	0.061	0.068	0.041	0.070

B

Decile of test score	Predicted probabilities as proportion of the sample mean				
	WLS	Talent	NLS-72	HSB	NELS
10	2.11	2.41	1.41	1.39	1.13
9	1.69	1.82	1.60	1.32	2.07
8	1.53	1.51	1.72	1.12	0.99
7	1.18	1.00	1.31	1.51	1.60
6	1.13	0.80	1.16	1.00	1.27
5	0.99	1.03	1.00	1.15	0.89
4	0.56	0.61	0.71	0.59	1.01
3	0.26	0.20	0.43	0.51	0.70
2	0.30	0.30	0.27	0.54	0.27
1	0.28	0.31	0.02	0.42	0.10

Notes: Values in panel A are the average predicted probability of entering the teaching profession, by decile, for a female with at least a high-school degree. Values in panel B are the predicted probabilities from panel A, normalized by the sample mean for each cohort.

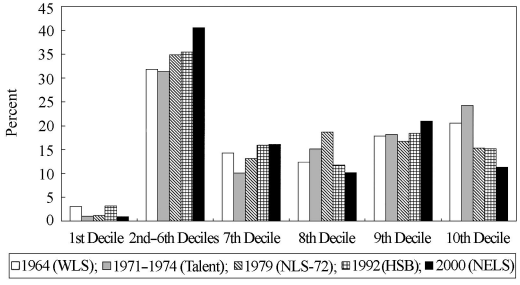

FIGURE 1. DISTRIBUTION OF TEACHERS ACROSS DECILE GROUPS, 1964–2000

percent for those in the top decile. While appropriate caution should be used with these results, given the small sample size, we find these results intriguing, and worthy of further study.

III. Conclusion

Despite a small number of cross-sectional studies that have examined the characteristics of college graduates choosing to enter teaching, 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 and in Corcoran et al. (2004), 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 (presumably those most likely to benefit from labor-market desegregation) became much less likely to enter the teaching profession than their peers near the middle of the distribution.

If our results can be applied to the wider population of new 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 was a student in 1964. For the casual observer, these results will surprise few.

REFERENCES

Bacolod, Marigee P. "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Supply and Teacher Quality 1940–1990." Mimeo, University of California–Irvine, 2003.

Ballou, Dale and Podgursky, Michael. "Teacher Pay and Teacher Quality." Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research, 1997.

Corcoran, Sean P.; Evans, William N. and Schwab, Robert M. "Women, the Labor Market, and the Declining Relative Quality of Teachers." *Journal of Policy Analysis and Management*, 2004 (forthcoming).

Ehrenberg, Ronald G. and Brewer, Dominic J. "Did Teachers' Verbal Ability and Race Matter in the 1960s? *Coleman* Revisited." *Economics of Education Review*, March 1995, *14*(1), pp. 1–21.

Ferguson, Ronald F. and Ladd, Helen F. "How and Why Money Matters: An Analysis of Alabama Schools" in Helen F. Ladd, ed., *Holding schools accountable: Performance-based reform in education.* Washington, DC: Brookings Institution, 1996, pp. 265–98.

Flyer, Frederick A. and Rosen, Sherwin. "The New Economics of Teachers and Education." *Journal of Labor Economics*, January 1997, *15*(1), pp. S104–39.

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

Hanushek, Eric A. and Pace, Richard R. "Who Chooses to Teach (and Why)?" *Economics of*

Education Review, June 1995, *14*(2), pp. 101–17.

Lakdawalla, Darius N. "The Declining Quality of Teachers." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8263, April 2001.

Levy, Harold O. "Why the Best Don't Teach." *New York Times*, 9 September 2000.

Loeb, Susanna and Page, Marianne E. "Examining the Link Between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-pecuniary Variation." *Review of Economics and Statistics*, August 2000, *82*(3), pp. 393–408.

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." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 1539, January 1985.

Murnane, R. J.; Singer, J. D.; Willett, J. B.; Kemple, J. J. and Olsen, R. J. *Who will teach?* Cambridge, MA: Harvard University Press, 1991.

Podgursky, Michael; Monroe, Ryan and Watson, Donald. "Teacher Mobility, Pay, and Academic Quality." Mimeo, University of Missouri, 2002; *Economics of Education Review*, 2004 (forthcoming).

Rice, Jennifer King. *Teacher quality: Understanding the effectiveness of teacher attributes.* Washington, DC: Economic Policy Institute, 2003.

Stoddard, Christiana. "Why Has the Number of Teachers per Student Risen While Teacher Quality Has Declined?" *Journal of Urban Economics*, May 2003, *53*(3), pp. 458–81.

Temin, Peter. "Teacher Quality and the Future of America." *Eastern Economic Journal*, 2002, *28*(3), pp. 285–300.

Wayne, Andrew J. and Youngs, Peter. "Teacher Characteristics and Student Achievement Gains: A Review." *Review of Educational Research*, Spring 2003, *73*(1), pp. 89–122.