Pulled Away or Pushed Out?

Explaining the Decline of Teacher Aptitude in the United States

CAROLINE M. HOXBY AND ANDREW LEIGH*

December 2003

Abstract

There are two main hypotheses for the decline in the aptitude of public school teachers since 1960: improved job opportunities for females in other occupations and the compression of teaching wages owing to unionization. Using data on several college graduating cohorts from 1961 to 1997, we investigate both hypotheses. To separate the hypotheses, we exploit the fact that states varied considerably in the progress of unionization and female wage parity. We proxy for a teacher's aptitude with the mean college aptitude of students at her undergraduate college. We identify the effects of unionization using laws that legalized and facilitated teachers' unionization. The evidence suggests that compression of teaching wages is responsible for about three-quarters of the decline in teacher aptitude. Females' opportunities in alternative occupations do matter, but opportunities improved rather similarly for females of all aptitudes. Although alternative occupations drew women out of teaching in general, they did not have a sufficiently disproportionate effect on high aptitude women to explain the bulk of the decline in teachers aptitude.

Authors' affiliations: Harvard University and NBER, NBER. Address: Department of Economics, Harvard University, Cambridge MA 02138. The authors gratefully acknowledge comments from Julian Betts and participants in the AEA session on Teacher Quality.

Logic suggests that a teacher's value-added is related to her academic aptitude. It is therefore troubling that teachers' aptitude has declined significantly in the United States since 1960, as demonstrated by Corcoran, Evans, and Schwab (2003). Combining longitudinal surveys, they find a marked fall in teachers' propensity to be in the top achievement quartile.

There are two main hypotheses for the decline in teacher aptitude. First, greater pay parity with males in non-teaching occupations may have drawn able women out of teaching. Second, unionization may have compressed pay, benefits, and non-monetary returns to aptitude in teaching, thereby pushed high aptitude people. In short, there is a "pull" hypothesis (pay parity in alternative opportunities) and a "push" hypothesis (pay compression in teaching). The hypotheses are not mutually exclusive. We need not choose between them; instead, we try to apportion blame.

To do this, we need variation in the timing and size of the pull and push factors. Fortunately, such variation exists among U.S. states. Our econometric identification relies on variation in labor markets and unionization laws among states at a point in time and within each state over time.

I. The Decision to Go into Teaching

The hypotheses about declining teacher aptitude can be exposited in a Roy model of occupational choice [Roy 1951]. If aptitude is positively correlated across occupations, the model predicts that compressing an occupation's pay-for-aptitude will push its high aptitude workers out. Increasing an occupation's pay across the board (for all aptitudes) will pull in workers from other occupations but will not necessarily change the distribution of aptitude between occupations. Only under restrictive conditions will an across-the-board increase in pay raise an occupation's mean aptitude.

Thus, if teachers' unionization compressed pay-for-aptitude, high aptitude people would migrate out. Similar migration would occur if non-teaching opportunities improved *disproportionately* for high aptitude women. However, if non-teaching opportunities improved similarly for female college graduates of all aptitudes, fewer of them would teach but teachers' aptitude would not necessarily decline.

II. Empirical Strategy

Start with a simple occupational choice equation:

(1) $$Prob\,(I_{ijt}^{tchr}=1) \;=\; \alpha_0 \;+\; \alpha_1 \ln(w_{ijt}^{f.tchr}) \;+\; \alpha_2 \ln(w_{ijt}^{f.alt}) \;+\; \boldsymbol{I_j^{state}}\,\boldsymbol{\alpha_3} \;+\; \boldsymbol{I_i^{apt}}\,\boldsymbol{\alpha_4} \;+\; \boldsymbol{I_t^{cohort}}\,\boldsymbol{\alpha_5} \;+\; \epsilon_{ijt}$$

The probability that female *i* from state *j* in cohort *t* teaches is a function of her pay in teaching ($\ln(w_{ijt}^{f.tchr})$), her pay in alternative jobs ($\ln(w_{ijt}^{f.alt})$), and other factors. We worry about other factors that are correlated with her state, aptitude, or cohort, so the equation includes state, aptitude, and cohort indicator variables.

We can decompose teacher pay into the deviation from the average pay in the state and the average pay in the state:

(2) $$\ln\left(w_{ijt}^{tchr}\right) \;=\; \ln\left(\frac{w_{ijt}^{f.tchr}}{\overline{w_{jt}^{f.tchr}}}\right) \;+\; \ln\left(\overline{w_{jt}^{f.tchr}}\right) \quad .$$

Unionization typically compresses variation in the first term and raises the second term.

We can decompose pay in alternative occupations into the part due to women's gaining pay parity with men and the part due to men's pay:

(3) $$\ln\left(w_{ijt}^{f.alt}\right) \;=\; \ln\left(\frac{w_{ijt}^{f.alt}}{w_{ijt}^{m.alt}}\right) + \ln\left(w_{ijt}^{m.alt}\right)$$

Thus equation (1) becomes:

(4) $$Prob\;(I_{ijt}^{tchr}=1) \;=\; \alpha_0 \;+\; \beta_1 \ln\left(\frac{w_{ijt}^{f.tchr}}{\overline{w_{jt}^{f.tchr}}}\right) \;+\; \beta_2 \ln\left(\overline{w_{jt}^{f.tchr}}\right) \;+$$
$$\beta_3 \ln\left(\frac{w_{ijt}^{f.alt}}{w_{ijt}^{m.alt}}\right) \;+\; \beta_4 \ln\left(w_{ijt}^{m.alt}\right) \;+\; \boldsymbol{I_j^{state}}\,\boldsymbol{\alpha_3} \;+\; \boldsymbol{I_j^{apt}}\,\boldsymbol{\alpha_4} \;+\; \boldsymbol{I_t^{cohort}}\,\boldsymbol{\alpha_5} \;+\; \epsilon_{ijt}$$

We cannot observe an individual female's pay as a teacher, in an alternative career, and as a male. We must therefore predict pay using people of the same cohort, state, and aptitude. Because we need to do this, we can, with negligible loss of information, estimate equation (4) with observations at the aptitude group-by-state-by-cohort cell level.

III. Identification

Because many factors can affect the pay of female teachers, we need instruments to isolate the effects of unionization. Our instruments are indicators for laws that facilitated or forestalled teachers' unionization. From 1955 onwards, some states enacted laws that gave teachers' organizations the rights to meet and confer with management, conduct collective bargaining, deduct members' dues and non-members' fees from paychecks, and exclude non-members from teaching. Other states enacted laws that protected non-members' right to work or prohibited paycheck deduction of dues and fees. Previous research has shown that the laws caused the speed and extent of teachers' unionization to vary, even among states with very similar labor markets such as Ohio and Illinois [Hoxby 1996, Saltzman 1988].

The instruments give us first stage equations for average teacher pay:

$$(5) \qquad \ln\left(\overline{w_{jt}^{f.tchr}}\right) = \delta_0 + \boldsymbol{UnionLaws_{jt}}\,\boldsymbol{\delta_1} + \boldsymbol{I_j^{state}}\,\boldsymbol{\delta_2} + \boldsymbol{I_t^{cohort}}\,\boldsymbol{\delta_3} + \epsilon_{jt},$$

and the ratio of pay in a given aptitude group to average pay:

$$(6) \qquad \ln\left(\frac{w_{ijt}^{f.tchr}}{\overline{w_{jt}^{f.tchr}}}\right) = \gamma_0 + \boldsymbol{UnionLaws_{jt}}\cdot\boldsymbol{I_i^{apt}}\,\boldsymbol{\gamma_1} + \boldsymbol{I_j^{state}}\,\boldsymbol{\gamma_2} + \boldsymbol{I_i^{apt}}\,\boldsymbol{\gamma_3} + \boldsymbol{I_t^{cohort}}\,\boldsymbol{\gamma_4} + \epsilon_{jt}.$$

The interaction terms in (6) allow unionization to have different effects for different aptitude groups–for instance, depressing the pay ratio for high aptitude females while raising it for others.

Pay parity is measured by the ratio of female-to-male earnings outside of teaching: $\ln\left(\frac{w_{ijt}^{f.alt}}{w_{ijt}^{m.alt}}\right)$. We believe that we do not need to instrument for this ratio because, *within* an aptitude group, changes in the ratio will not merely reflect changes in the aptitude of women who choose to work. Also, we do not expect the ratio to be endogenous to events in teaching. It is possible, for instance, that unionization drove high aptitude women into law and thereby raised the pay of female attorneys. However, such phenomena were probably unimportant. Not instrumenting will, if anything, make us overstate the role of pay parity. Keep this in mind when interpreting the results.

Male pay in non-teaching occupations ($\ln\left(w_{ijt}^{m.alt}\right)$) is probably correlated with unobserved changes in a state's economy, technology, and culture. Because we do not have an instrument for this variable, it is fortunate that we do not need its structural coefficient to test the hypotheses. We include it as a control but discourage literal interpretation of its coefficient.

IV. Data

We need earnings and occupation data that are linked to a measure of aptitude, cover most states, are comparable over time, and have good coverage of college graduates (over the period of interest, only college graduates become teachers). We use the surveys of Recent College Graduates (RCG), which cover the baccalaureate classes of 1975, 1977, 1980, 1984, 1986, and 1990; two predecessors of RCG that cover the classes of 1961 and 1964 to 1967; and two successors to RCG that cover the classes of 1993 and 1997. Data details are in an appendix (www.nber.org/~hoxbyleigh).

We record occupation and pay 18 months to two years after the baccalaureate degree. Lacking a direct measure of aptitude, we link people to the mean combined SAT scores of their college and then divide them into six groups: those from colleges with SAT scores in the top five percentiles, the next ten, the next fifteen, the next twenty, the next 25, and the bottom 25 percentiles. The SAT cut-points are constant over time so that aptitude is defined in absolute terms. The aptitude groups are finer at the top of the distribution because previous research suggests that the top quartile accounts disproportionately for the decline in teacher aptitude.

We aggregate data to the aptitude group-by-state-by-cohort cell. For instance, the dependent variable in our regression is the share of female college graduates in an aptitude group in a state in a cohort who become public school teachers.

V. Push and Pull Factors and the Decision to Become a Teacher

Table 1 shows changes in the earnings variables from 1963 to 2000 (the first and last years of our earnings data). The earnings of the average female teacher rose by 8 percent in real terms from 1963 to 2000.

Row (2) shows that the ratio of the lowest aptitude teachers' earnings to mean teacher earnings rose 0.33 natural log points from 1963 to 2000. For instance, if they began with an earnings ratio of 0.72, they ended with a ratio of one (parity with the mean teacher's earnings). The ratio of the highest aptitude teachers' earnings to mean teacher earnings fell 0.45 natural log points over the same period. If they began with an earnings ratio of 1.59, they ended with a ratio of one. By 2000, most states had earnings ratios near one *for all aptitude groups*.

Row (3) shows that, for a college graduate in one of the top three aptitude groups, the ratio of female to male earnings in non-teaching occupations rose 0.08 to 0.10 natural log points. She could expect the ratio of her earnings to similar aptitude males' earnings to rise from 0.77 in 1963 to 0.86 in 2000. For a lower aptitude woman, there was little change in the female to male earnings ratio (the small decline for the lowest group is an artifact of the group's having a poorly defined aptitude floor). She could expect the ratio of her earnings to similar aptitude males' earnings to hover around 0.81.

Row (4) shows that the real earnings in non-teaching occupations rose by about 33 percent for most college graduate men, but by 42 percent for men in the highest aptitude group.

Rows (5) and (6) show that the share of lowest aptitude female college graduates who became teachers fell from 48 to 16 percent between 1963 and 2000. Over the same period, the share of highest aptitude female college graduates who became teachers fell from 20 to 4 percent. Rows (5) and (6) show the dependent variable we use in our regression. However, the groups differ in size so we must weight them to compute the overall effect on teacher aptitude. Rows (7) and (8) show that, between 1963 and

2000, the share of all teachers who came from the lowest aptitude group rose from 16 to 36 percent, while the share from the highest aptitude group fell from 5 to 1 percent.

VI. Why Teacher Aptitude Declined

Table 2 presents instrumental variables estimates of equation (4). The first coefficient shows that the higher was the ratio of teacher earnings for one's aptitude group to mean teacher earnings, the more likely one was to teach. The ratio rose by 0.33 natural log points for the lowest aptitude group and fell by 0.45 natural log points for the highest aptitude group. Therefore, pay compression increased the share of the lowest aptitude female college graduates who became teachers by about 9 percentage points and decreased the share of the highest aptitude female college graduates who become teachers by about 12 percentage points.

The second coefficient indicates that the share of female college graduates who taught rose by about 2 percentage points as a result of the observed 8 percent increase in the real earnings of the average teacher. Conveniently, an 8 percent increase in earnings is approximately the effect of unionization [Hoxby, 1996]. These estimates cannot explain much of the decline in teacher aptitude because women from all groups *necessarily* experience the same increase in mean pay. (Teacher aptitude declines slightly because the same percentage point increase is applied to aptitude groups of different size, with the lower groups being larger).

The third coefficient indicates that the higher is the ratio of female to male earnings in non-teaching occupations, the less likely women are to teach. Specifically, improvements in pay parity decreased the fraction of women who taught by 3.2 percentage points for the highest aptitude group, 2.5 percentage points for the top three aptitude groups, and zero for the three lower aptitude groups.

The coefficient on the pay parity variable is roughly a mirror image of the coefficients on the teaching pay variables, suggesting that teaching and non-teaching pay similarly affect the decision to teach. However, pay parity in alternative occupations explains much less of the decline in teacher aptitude than does the compression of teachers' pay. Why is this? Although the coefficients are similar, the difference between high and low aptitude women's experience is much smaller for pay parity than for pay compression. Female-male parity improved pay 0.13 log points more for the highest aptitude women than the lowest aptitude women. However, compression worsened pay 0.78 log points more for the highest aptitude women than the lowest aptitude women. The difference in the *x*'s drives everything: any plausible coefficient estimates would suggest a large role for pay compression, relative to pay parity.

The fourth coefficient indicates that factors correlated with male earnings reduced the share of women who became teachers by 11 percentage points for the highest aptitude group and by 8 percentage points for the other groups. These hard-to-interpret results show that factors correlated with male earnings only help to explain the decline in teaching among the highest aptitude group. Applying the same percentage point increase to the other five groups would actually raise teacher aptitude slightly because the lower aptitude groups are larger.

It is not surprising that the coefficients on the pay variables are similar in absolute value. Given the decomposition of equation (1), we expect similarity between β_1 and β_2 and between β_3 and β_4. The occupational choice model makes us expect similarity among all four coefficients.

VII. Apportioning "Blame"

To apportion blame for the decline in teacher aptitude, we apply the estimated coefficients to the changes in the earnings variables, taking account of the aptitude group sizes (important). Such computations allow us to say, for instance, how many teachers would have high aptitude if pay compression or pay parity had not changed.

The share of teachers in the highest aptitude category fell from 5 percent to 1 percent. Of this change, pay compression explains about 80 percent, pay parity explains about 9 percent, and the change in mean teacher earnings explains about 1 percent. If we accept the coefficient on male earnings at face value, it explains another 19 percent, but this almost certainly overstates the causal effect.

The share of teachers in the lowest aptitude category rose from 16 to 36 percent. Of this change, pay compression explains about 25 percent, pay parity explains about 6 percent, the change in mean teacher earnings explains about 2 percent, and (if we accept the coefficient at face value) male earnings explain another 8 percent. The remainder is explained by the increase in the size of the lowest aptitude group–the number of women with this aptitude who graduate from college.

When we began this study, our prior was that pay parity would play the major and pay compression the minor role. We had not recognized the implications of the fact that pay parity changed similarly for college women of all aptitudes, which makes its smaller role predictable. Put another way, outside of teaching, high aptitude college women did not gain dramatically relative to low aptitude college women: they all gained over time. However, in teaching, high aptitude women experienced substantial relative losses.

References

Corcoran, Sean; Evans, William and Robert Schwab. "Changing Labor Market Opportunities for Women and the Quality of Teachers 1957-1992." National Bureau of Economic Research Working Paper No. 9180, 2002.

Hoxby, Caroline M. "How Teachers' Unions Affect Education Production," Quarterly Journal of Economics, August 1996, 111(3), pp. 671-718.

Roy, A.D. "Some Thoughts on the Distribution of Earnings." Oxford Economic Papers, June 1951, 3, pp. 135-46.

Saltzman, Gregory. "Public Sector Bargaining Laws Really Matter: Evidence from Ohio and Illinois" in Richard Freeman and Casey Ichniowski, eds., When Public Sector Workers Unionize. Chicago: University of Chicago Press, 1988, pp. 41-80.

Table 1

		Aptitude Category					
		lowest	2	3	4	5	highest
(1)	$\Delta\ \ln(\overline{w_{jt}^{f.tchr}})$ from 1963 to 2000	0.08 (same for all)					
(2)	$\Delta\ \ln(w_{ijt}^{f.tchr} / \overline{w_{jt}^{f.tchr}})$ from 1963 to 2000	0.33	0.29	0.08	-0.14	-0.32	-0.45
(3)	$\Delta\ln(w_{ijt}^{f.alt} / w_{ijt}^{m.alt})$ from 1963 to 2000	-0.04	-0.01	0.01	0.10	0.08	0.09
(4)	$\Delta\ln(w_{ijt}^{m.alt})$ from 1963 to 2000	0.31	0.30	0.31	0.29	0.33	0.42
(5)	share *who are* tchrs 1963	0.48	0.41	0.41	0.34	0.30	0.20
(6)	" " 2000	0.16	0.18	0.13	0.12	0.10	0.04
(7)	share *of* all tchrs in 1963	0.16	0.26	0.26	0.15	0.12	0.05
(8)	" " 2000	0.36	0.31	0.17	0.11	0.04	0.01

Table 2

dependent variable: share who teach in public elementary or secondary school	
$\ln(w_{ijt}^{f.tchr} / \overline{w_{jt}^{f.tchr}})^{\dagger}$	0.27** (0.11)
$\ln(\overline{w_{jt}^{f.tchr}})^{\dagger}$	0.29** (0.13)
$\ln(w_{ijt}^{f.alt} / w_{ijt}^{m.alt})$	-0.25** (0.09)
$\ln(w_{ijt}^{m.alt})$	-0.26* (0.08)
state, aptitude, and cohort (time) fixed effects	yes
instruments (excluded from second stage)	union laws
F-stat (Prob>F) from first-stage: jt test on excluded instruments	1.88 (0.0007)

Notes: Instrumental variables regression using 1326 observations at the aptitude group-by-state-by-cohort level. † indicates that the variable is treated as endogenous. Standard errors are in parentheses. ** (*) indicates that the coefficient is statistically significantly different from zero with 95 percent (90 percent) confidence.