Divergent Trends in U.S. Teacher Quality: 1980-2010

Richey, Jeremiah

Kyungpook National University

30 April 2014

Divergent Trends in U.S. Teacher Quality 1980-2010

Jeremiah Richey
jarichey80@gmail.com

Kyungpook National University
April 30, 2014

Abstract

This paper documents the changes in the ability distribution of individuals entering the teaching profession among individuals born between 1957-1964 and 1980-1984. This is done using the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth and a constructed Armed Force Qualifying Test score that allows direct comparison of ability between cohorts. Such direct comparisons between cohorts was previously not possible due to a lack of directly comparable measures of ability. I find there are minimal differences in ability between cohorts with the only significant difference being a small drop in ability around the 75th percentile of the conditional distribution. However, this similarity across cohorts masks vast differences in specific demographics. In particular, white women are of much lower ability in the 1997 cohort whereas minority women are of substantially higher ability. Moreover the gain in ability of female minority teachers is due to both a cohort improvement in ability and an increase in the within cohort ability of teachers. Puzzling is the significant negative trend in ability in the 1979 cohort which is not apparent between cohorts. This implies a likely 'bottoming out' and rebound in teacher ability in the 1990s.

JEL Classification:

Key Words: Education, Teacher Quality

1 Introduction

The education profession, a profession historically dominated by women, has faced two great head winds over the last several decades. One is the continued desegregation of the U.S. labor market and parallel rise of women in the labor force. The other is continued technological advances that have markedly increased the marginal product of labor in most industries outside of education while leaving the education industry largely untouched. Both of these trends have affected the profession's relation with its historic labor pool. For a sense of this, note the female labor force participation rate increased nearly 20% from 1980 to the turn of the century with the percentage of women in high-paying professional and managerial jobs seeing great gains (Blau and Kahn 2007). Simultaneously the female to male earnings ratio substantially increased from 60% in 1980 to 77% in 2010 (Institute for Women's Policy Research 2014). Meanwhile the U.S. average annual public teacher salary increased by only about 15% in real terms over 25 years from 1985 to 2010 (National Education Association 2010) during which time the average woman's yearly income increased over 50%.

An interesting and important question then is what consequences the meeting of these trends has had for the ability of individuals entering the teaching profession. If women are increasingly likely to work in industries outside of education where their real wages have been steadily increasing, while the wages in education have been relatively stagnant, one quickly comes to the hypothesis that high ability women with increased outside options will be less likely to become teachers leading to detrimental outcomes for teacher quality. And this is likely a public concern as teacher cognitive ability has been linked to student achievement (see Rice (2003) for a survey). However research in this area appears limited in scope[1]. Corcoran, Evans and Schwab (2004 a,b) provide perhaps the most concise documentation of this trend. They compare five cohorts who graduated from high school between 1957 and 1992. Their main finding is that the probability of women in the highest deciles of within-cohort measured ability entering the teaching profession seems to have declined. Conversely women from lower deciles have become more likely to enter the profession.

[1]Recent exceptions are Bacolod (2007) and Corcoran, Evans and Schwab (2004a,b), Hoxby and Leigh (2004), Lakdawalla (2001), and Flyer and Rosen (1997).

However, these findings are limited in the information they reveal regarding the change in the ability distribution of teachers - it does not allow for comparisons of ability across cohorts (a fact they note). Bacolod similarly notes the decrease in the probability of female teachers scoring in the upper deciles of standardized tests becoming teachers over cohorts born 1941-1964. However her focus and main contribution is linking relative wages to probability of becoming a teacher, specifically how professional-to-teacher relative wages have large impacts on high ability women's choice of profession. Others have linked declines in teacher quality to technological changes in the U.S. labor market using relative wages as proxy for teacher ability (Lakdawalla 2001) and to increases in teacher unionization using undergraduate institution-level SAT score as a proxy for ability (Hoxby, Minter and Leight (2004)).

This paper adds to the literature on teacher quality by documenting the changes in the ability distribution of young people entering the teaching profession over a 20 year period between the 80's and the 2000's. This is done using the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth (NLSY). Importantly the measure of ability is a constructed Armed Force Qualifying Test (AFQT) score that allows direct comparison of ability across cohorts - something previous research has been unable to do. I not only document the change in the ability distribution of teachers, but also the change in the within-cohort relative ability of teachers in a directly comparable way. Briefly, I find the distribution of ability has changed little between cohorts, but this masks starkly different trends for different demographic groups. White female teachers present large drops in ability, minority female teachers present large gains in ability, and men present no noticeable change in ability. Puzzling though, there appears a strong negative trend in teacher ability within the 1979 cohort implying a likely 'bottoming out' in teacher ability in the 1990's from which there was a rebound. With regards to the changes in the relative ability of teachers, the results are qualitatively similar with small or insignificant cohort effects masking divergent trends in specific demographic groups. Furthermore, comparisons of the changes in absolute and relative ability of teachers highlights the benefit of using directly comparable test scores. The rest of the paper is organized as follows. Section 2 discusses the data. Section 3 presents the main findings and Section 4 discusses the main results and concludes.

2 Data

The data used comes from the 1979 and 1997 NLSY. The 1979 NLSY is survey of youths aged 14-22 in 1979. It is comprised of a cross-sectional survey representative of the nation (n = 6,111), an over sample of minorities and poor whites (n = 5,295) and a sample of the military (n = 1,280) . The 1997 NLSY is survey of youths aged 12-18 in 1997. It is comprised of a cross-sectional survey representative of the nation (n = 6,748) and an over sample of minorities (n = 2,236). The sample used for this survey is the cross-sectional survey and over sample of minorities for both cohorts (the over sample of the military poor whites from the 1979 cohort are excluded and were discontinued in 1984 and 1990).

The main variables of interest are reported occupation and a measure of ability. Occupation is characterised as whether one has reported being a teacher between the ages of 21 and 31[2][3]. The measure of ability is test scores from the AFQT. An important issue when comparing a measure of ability across individuals is comparability of the available measure. Differing test instruments have prevented previous researchers from directly comparing the ability of teachers across time. The following describes the construction of AFQT scores that are directly comparable across cohorts.

The Armed Services Vocational Aptitude Battery (ASVAB) was administered to both the 1979 and 1997 cohorts of the NLSY from which AFQT scores are constructed. However these scores are not directly comparable as the two cohorts took different versions of the ASVAB. The 97 cohort took a computer administered test (CAT) while the 79 cohort took a paper and pencil (P&P) version. Additionally the test was administered at different ages for the two cohorts. A two step process is taken to make the two scores directly compatible and is documented in greater detail by Altonji, Bharadwaj and Lange (2009). First, a mapping from the P&P to the CAT version of the test is used to make the raw scores equivalent. This mapping is due to work by Segal (1997) and based on a sample of individuals randomly assigned the P&P or CAT version between 1988 and 1992.

[2]In the latest wave of the 1997 NLSY available the oldest respondents are 31.

[3]Teacher is defined as pre-kindergarten, kindergarten, elementary, or secondary school teacher based on 1990 and 2002 Census occupation codes. However for first 3 years of the 1979 NLSY only 1970 census occupation codes are available and so teacher is defined as elementary or secondary school teacher. When re-analyzed excluding pre-K and kindergarten teachers the results are qualitatively identical and very similar quantitatively as well.

Then a equi-percentile mapping is used across age groups to create age-consistent scores (Altonji, Bharadwaj and Lange (2009)). This is done by mapping both cohorts into their cohort specific 16 year old score distribution (age 16 is the age group with the greatest overlap between the two cohorts). The result is directly comparable AFQT scores[4].

3 Results

3.1 Shift in Teacher Ability Across Cohorts

Having directly comparable test scores for both cohorts allows direct comparisons of the ability distribution of teachers between and within the two cohorts. This is something which previous authors were unable to do. Table 1 compares the mean and select quantiles of the two cohort distributions for demographic groups of interest while Figure 1 (a) - (c) graphs the cohort effect across the entire quantile range (these and all regressions use sample weights to correct for the sampling method of the NLSY). When the focus is the entire set of individuals reporting teaching at least one year the mean difference is slightly lower for the 1997 cohort yet this difference is insignificant. When looking at the quantiles the cohort effect grows negative across the quantiles yet is only significant around the 75th percentile where there is a loss of about 3.5 points compared to a score of 200 for the 1979 cohort at this percentile. Tests of a differing cohort effect for men and women indicate women have a slightly more negative effect in mean and across quantiles, yet these differences are not significant. However, when focusing on women, there are significant differences in cohort effects between white and minority women. As such each of these sub-populations is investigated separately.

For white women the cohort effect leads to a drop in mean AFQT score of almost 6.5 points compared to the 1979 average of 185. This effect is somewhat stronger (point estimate) at the lower quantiles, yet fairly stable throughout most of the conditional distribution at about 6 points

[4]Altonji, Bharadwaj and Lange have very kindly made these constructed scores publicly available on Fabian Lange's website http://www.econ.yale.edu/ fl88/ and I thank them for their generosity in this regard.

before converging to a zero effect beginning around the 85th percentile. Minority women on the other hand have a strong positive cohort effect. The mean AFQT score for the 1997 cohort is nearly 20 points higher than the mean score of about 145 for the 1979 cohort. Moreover this effect has a distinct inverted U shape (which can be seen in Figure 1 (c)) with the cohort effect being strongest around the 60th percentile where it peaks at around 30 points. Given that minority women have lower scores than white women, this increase in minority women's scores helps to offset the fall in white women's, but only in the lower end of the distribution, thus a decline in ability still appears in the upper quantiles for the full distribution.

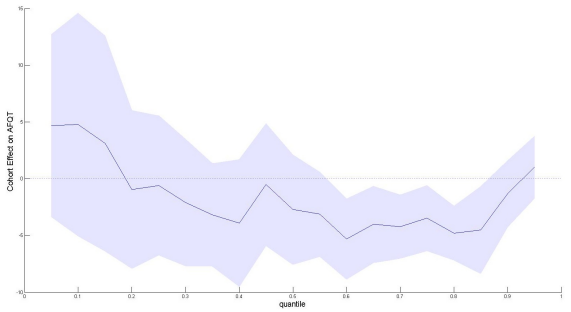
Helpful in understanding this cohort effect on the ability of teachers is to visualize the actual shift in the ability distribution. These distributional shifts are graphed in Figure 1 with panels (d) - (f) graphing the CDFs and panels (g) - (i) plotting the PDFs for the 1979 and 1997 cohorts. This allows one to actually see the negative but small shift in the entire teacher population, the much larger negative shift in the white female population, and the stunning change in the distribution of minority female teachers.

A large concern facing the teaching profession is the high degree of attrition (see Borman and Dowling (2008) for an overview). With this is mind, the above analysis is rerun on the sub-population of teachers reporting working at least two years as a teacher. These results are given in second set of columns in Table 1. Qualitatively the relationships do not change with this restriction, however the effects are somewhat magnified for the full population and white women yet subdued for minority women. The mean effect when using the entire sample increases to 3.4 point loss between cohorts and this effect becomes significant, and again similar trends emerge across the quantiles with the largest negative effect appearing around the 75 percentile. Additionally I reanalyze the data where the weights are adjusted to reflect the number of years an individual reported being a teacher (placing greater weight on those working more years) and again the results are nearly identical to those for the reduced population of teachers working 2+ years.

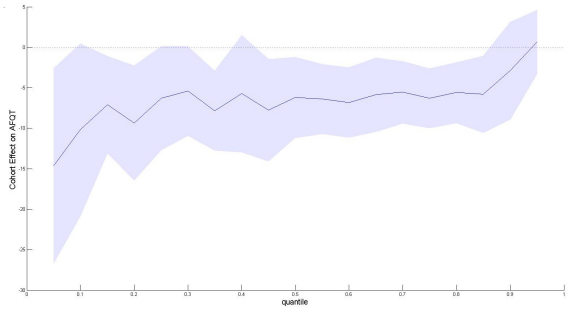
Table 1: Results: Cohort Effect on Teacher Ability. Dependent Variable is Constructed AFQT Score.

		Quantile						Quantile				
Dep. Variable	OLS	10%	25%	50%	75%	90%	OLS	10%	25%	50%	75%	90%
	Full Sample: n = 898						*Full Sample 2 yrs +: n = 652*					
Intercept	180.27	141.86	165.52	186.20	199.69	208.00	182.52	151.59	169.50	187.30	201.37	208.00
	(1.18)	(3.84)	(1.86)	(2.04)	(0.98)	(0.87)	(1.36)	(4.87)	(2.59)	(2.02)	(1.28)	(0.92)
1997	-1.83	4.76	-0.59	-2.71	-3.47	-1.31	-3.40	-3.98	-3.21	-3.82	-5.15	-1.31
Cohort	(1.65)	(4.47)	(2.96)	(2.34)	(1.59)	(1.79)	(1.71)	(5.50)	(3.02)	(2.52)	(1.62)	(2.00)
	White Female: n = 416						*White Female 2 yrs +: n = 312*					
Intercept	185.69	162.60	173.69	188.99	200.38	208.00	187.74	164.69	174.62	191.72	201.37	208.00
	(1.73)	(4.76)	(2.77)	(2.27)	(1.52)	(1.34)	(1.34)	(1.97)	(2.50)	(2.28)	(1.44)	(1.39)
1997	-6.34	-10.17	-6.30	-6.18	-6.29	-2.88	-7.52	-10.88	-7.23	-8.23	-7.28	-2.88
Cohort	(2.12)	(5.84)	(3.63)	(2.75)	(1.81)	(2.74)	(1.94)	(3.39)	(3.06)	(2.82)	(2.13)	(2.84)
	Minority Female: n = 301						*Minority Female 2 yrs +: n = 216*					
Intercept	144.87	116.30	125.89	141.86	160.81	179.57	147.81	118.42	125.89	143.91	166.55	184.27
	(2.00)	(1.67)	(1.63)	(2.07)	(3.34)	(5.05)	(2.12)	(2.36)	(2.35)	(4.25)	(3.25)	(3.81)
1997	19.30	8.54	21.73	27.24	25.96	14.52	15.86	6.41	20.73	24.62	20.23	10.65
Cohort	(3.31)	(8.13)	(3.44)	(3.78)	(3.54)	(5.53)	(3.33)	(7.92)	(3.93)	(5.24)	(3.58)	(4.17)

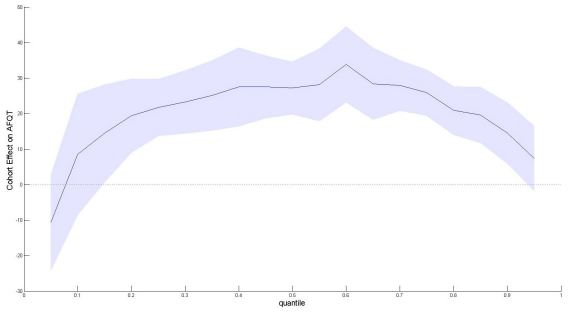
Bootsrapped SEs are reported in parenthesis



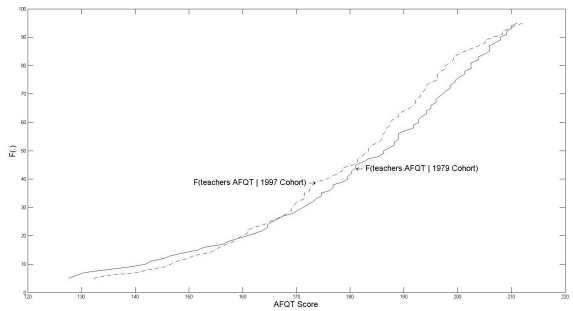
(a) Full Sample Effect



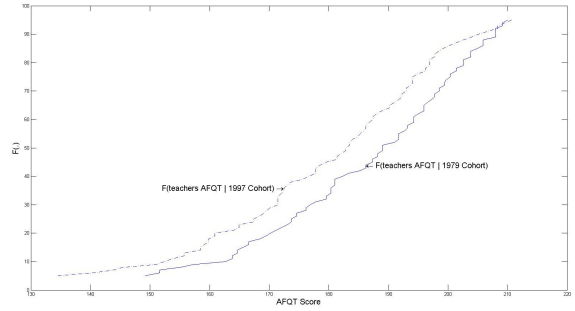
(b) White Women Effect



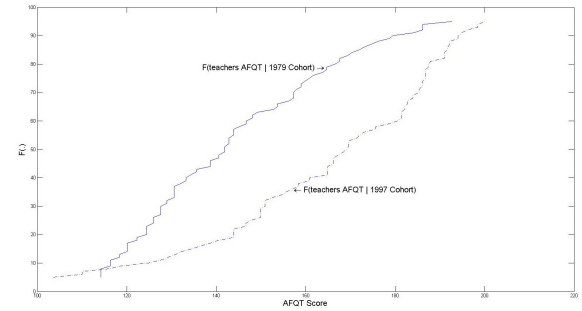
(c) Minority Women Effect



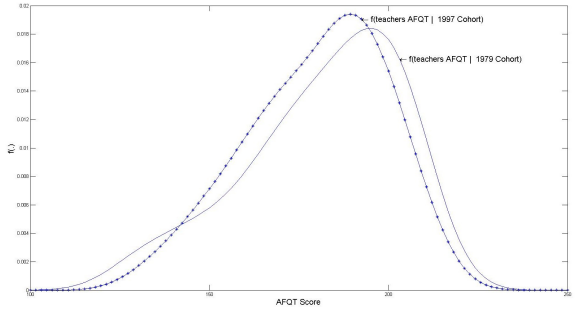
(d) Full Sample CDF



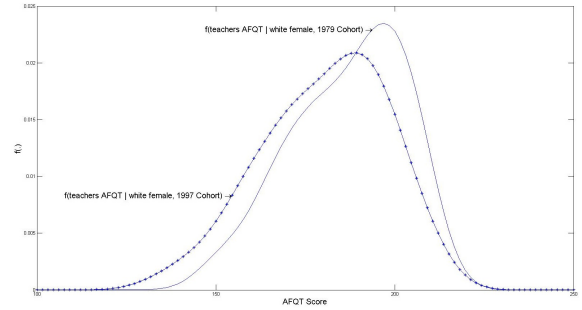
(e) White Women CDF



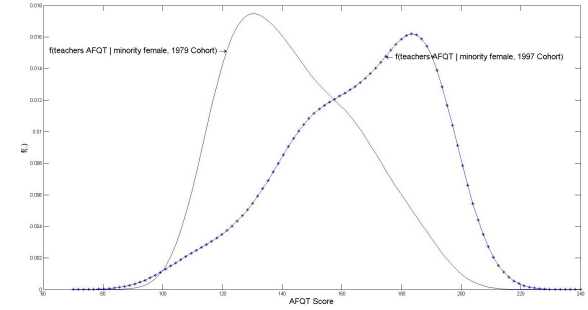
(f) Minority Women CDF



(g) Full Sample PDF



(h) White Women PDF



(i) Minority Women PDF

Figure 1: Cohort Effect on AFQT Score for Teachers and Implied Shift in Distribution.

3.2 Within Cohort Trends

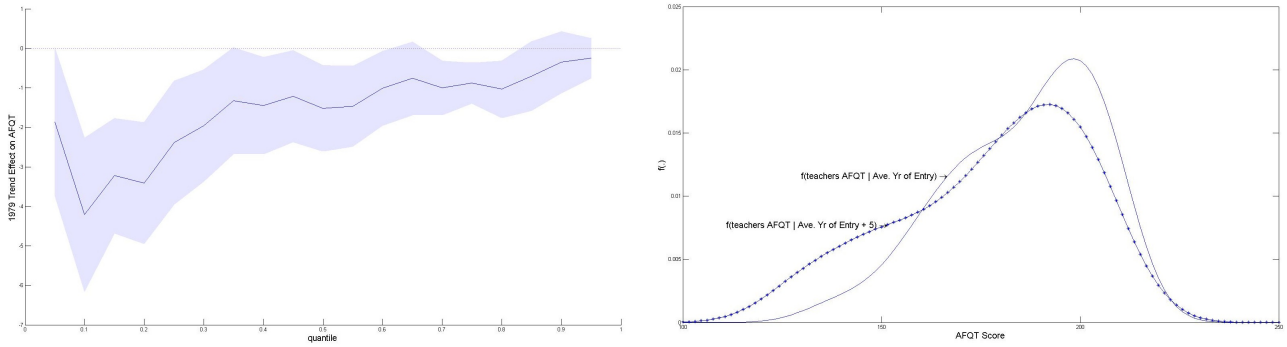
The above analysis treated the 1997 and 1979 cohorts as single entities separated by about 20 years (the average teacher in the 79 cohort began teaching in 1987 and the average teacher in the 97 cohort began teaching in 2006), when in reality the 79 cohort spans 8 birth years (1957-1964) covering individuals entering the profession between 1982 and 1995 while the 97 cohort spans 5 birth years (1980-1984) and covers individuals entering the profession between 2001 and 2010. An interesting question then is if there are any within-cohort trends in teacher quality. To investigate this I regress AFQT scores for teachers on a year trend for when the individual first reported teaching. When looking at the 1979 cohort there is a clear trend of within cohort decline both in the mean and in the various quantiles with the largest effects being in the lower quantiles (when run on the reduces population of those reporting teaching 2 or more years the results are qualitatively similar though with a reduced magnitude). Moreover there does not appear to be any differences within cohort between men and women or minority and white women with regards to the decline (this result comes from unreported results with interaction terms, though when run separately men have a negative but insignificant effect). On the other hand when looking at the 1997 cohort there is no significant trend within cohort (again unreported results with interaction effects imply that there are significant cohort differences in trend though). Neither is there any sign of a difference in the lack of trend between men and women. However, again there is a difference between white and minority women in their within-cohort trend. Over the 1997 cohort white women recover somewhat in the lower end of the distribution. Conversely, scores for minority women exhibit a small negative trend over the cohort years.

Again here it helps to visualize this trend effect graphically. Panel (a) in Figure 2 plots the trend effect across quantiles. Perhaps most interesting however is the estimated effect on the ability distribution of the trend and is plotted in panel (b) of Figure 2. This shows the expected shift in the ability distribution had the estimated trend effect in the 1979 cohort continued for 5 years. When comparing this figure with the shift between cohorts seen in panel (h) of Figure 1 it becomes apparent that there appeared to be a trend over the 1979 cohort (entered teaching profession 1982-1995) that is not apparent between cohorts. This implies there was a halting and reversal of this

Table 2: Results: Dependent Variable is Standardized AFQT Score.

		Quantile				
Dep. Variable	OLS	10%	25%	50%	75%	90%
		Full Sample 1979				
Intercept	191.67	176.40	182.56	197.91	205.76	210.82
	(3.34)	(8.17)	(4.49)	(5.79)	(2.17)	(2.56)
Year	-1.51	-4.21	-2.38	-1.52	-0.88	-0.35
Trend	(0.43)	(0.89)	(0.82)	(0.68)	(0.27)	(0.36)

Bootsrapped SEs are reported in parenthesis



(a) Effect

(b) Effect of 5 Yr Shift

Figure 2: Trend Effect on AFQT Distribution of Teachers Within 79 Cohort

negative trend in ability (especially in the lower end of the ability distribution). Again due to concerns over attrition in the teaching industry I rerun the above analysis with those teaching at least two years. For the full 1979 cohort the negative trend still emerges though at a subdued rate, particulary in the lower quantiles, with the mean effect only slightly over half the effect in Table 4. Again no trend emerges in the 1997 cohort for any population.

3.3 Teacher vs. Non-Teacher

A different, yet related question regards the ability of teachers in comparison to their cohort and how this relative ability may have changed over time. This was the sole avenue of comparisons available to previous authors investigated the trends in teacher ability. Here I similarly report changes in within-cohort teacher standing. This allows the previous results on changes in teacher

ability to be decomposed into cohort wide changes and relative within-cohort changes. Again the availability of a directly comparable scores allows greater transparency in these changes over relative measures in terms of standard deviations (as in Corcoran, Evans and Schwab 2004 a) which can be sensitive changes in the second moment of the ability distribution. This is done by regressing standardized AFQT score on a dummy for 1997 cohort, a dummy for reporting being a teacher, and an interaction term (along with intercept). The key item of interest then is the coefficient on the interaction term and also the ratio of it to the coefficient on the occupation dummy[5]. Results are given in Table 3 and are again given for the mean and key quantiles. Results from a standard regression find a small and insignificant effect for the interaction term, implying on average there has been no change in relative ability of those entering teaching. However there is a significant effect at the 50th percentile. The coefficient on the interaction term is -4.67 which is about 33% of the coefficient on the teacher dummy at this percentile. A standard Chow test rejects the null of all of the coefficients being similar for men and women and similarly rejects the coefficients being identical for minority and white women. As such I present results parallel to those in Table 1 reporting effects on the entire sample, those for white women, and those for minority women (again the effect on men is not significantly different from zero in any trends though there is a marginally significant loss in the teacher gap for men between cohorts at the median).

The results here for white women are again in stark contrast to those for the entire sample. The cohort-teacher interaction effect is measured at -7.88 which is nearly 60% of the teacher effect for white women. This means that, on average, white female teachers have lost over half of their ability gap compared to non-teachers between years of the 1979 and 1997 cohort. The quantile results reveal more trends in this effect. At the 75th percentile, these teachers have lost nearly 100% of their ability gap, 66% at the 50th percentile, and 50% at the 25th percentile. This result is parallel to previous findings in that the loss in teacher ability for women seems strongest in the upper percentiles of the distribution.

As in section 3.1, a striking trend is again revealed when the focus is minority women. The results

[5]There are well known issues with inference regarding ratios of coefficients. This issue is commonly termed 'Fieller's Problem'. While these concerns are noted, I do not directly address these inferential issues and caution the reader that my discussion of the ratios of these coefficients should be interpreted with this in mind.

Table 3: Results: Dependent Variable is Standardized AFQT Score.

		Quantile				
Dep. Variable	OLS	10%	25%	50%	75%	90%
		Full Sample				
1997 Cohort	0.98	-1.35	4.76	1.96	-1.58	0.08
	(0.52)	(1.34)	(0.81)	(0.71)	(0.68)	(0.81)
Teacher	15.11	21.76	23.66	16.70	7.97	4.17
	(1.41)	(4.06)	(1.92)	(1.70)	(1.27)	(0.86)
Cohort *	-2.81	6.11	-5.35	-4.67	-1.89	-1.39
Teacher	(2.06)	(5.25)	(3.02)	(2.19)	(1.78)	(1.95)
		White Women				
1997 Cohort	1.54	0.85	3.43	3.16	0.46	1.69
	(0.92)	(2.80)	(1.29)	(1.38)	(0.84)	(1.50)
Teacher	13.25	27.02	18.69	14.37	7.13	3.00
	(1.83)	(4.51)	(2.37)	(1.96)	(1.31)	(1.39)
Cohort *	-7.88	-11.03	-9.73	-9.35	-6.75	-4.56
Teacher	(2.78)	(6.19)	(3.28)	(2.99)	(2.00)	(3.67)
		Minority Women				
1997 Cohort	12.12	6.00	7.43	16.61	16.31	15.51
	(0.94)	(2.07)	(1.65)	(1.37)	(1.37)	(1.85)
Teacher	5.78	12.23	5.79	4.66	3.48	4.94
	(2.39)	(2.20)	(2.06)	(2.51)	(3.30)	(4.49)
Cohort *	7.19	2.54	14.29	10.63	9.66	-0.99
Teacher	(3.38)	(8.23)	(3.60)	(3.81)	(3.74)	(5.41)

For OLS robust SEs are in parenthesis. For quantiles, bootstrapped SEs are in parenthesis.

for a standard mean regression in Table 5 show a strong positive cohort-teacher interaction term emerges. In fact the estimated effect is 7.19 - nearly 125% of the teacher effect. This implies that minority female teachers have more than doubled their average ability gap over their comparative non-teaching population. Similar strong results hold across the quantiles. At the 75th percentile the cohort-teacher effect is 9.66 which is 275% of the estimated teacher effect, and the relation at the 50th and 25th percentiles are over 200% and nearly 300%. This implies minority female

teachers are of a significantly higher ability in the 1997 cohort than the 1979 cohort because of not only a cohort wide effect but also a significant increase in relative ability in those who choose the teaching profession.

Comparing these results to the cohort comparison in section 3.1 reveals the benefit of using directly comparable AFQT scores for a measure of ability. If the scores were not directly comparable the only measure of change in teacher ability would be changes in a measure of relative ability, and this would omit the cohort wide effects that exasperate or mitigate these changes. For example, the decline in overall teacher ability, as well as that for white women, would be overestimated, while the gain in teacher ability for minority women would be underestimated.

In both absolute and relative measures there appears to be substantial losses in the ability of white female teachers between cohorts. However this is substantially mitigated in the full teacher population by strong gains in ability in minority women and relatively constant levels of ability in men. It is worth noting then that the (weighted) percentage of minority women teachers in the NLSY is somewhat larger then found in Census data. However the percentage increase between the cohorts is similar[6]. Also the percentage of male teachers is larger in the NLSY (than Census or ACS data) and exhibits larger gains in participation. However this should not be too large a concern as males are not the driving force behind the overall change in ability of teachers which is more a result of the interaction of the female populations.

4 Summary and Conclusion

This paper has documented the change in ability of individuals entering the teaching profession by comparing directly comparable AFQT scores for the 1979 and 1997 cohorts of the NLSY. Such direct comparison of ability across cohorts of teachers was previously infeasible due to data limitations. Several key findings emerge.

[6]In the NLSY the (weighted) percentage of teachers who are minority women increased from 15 to 17.8% between cohorts. Compared to Census data for 1990/2000/2010 for which (of comparable age groups) they composed 10.5, 13 and 10.5% with consistent 11-12% in the earlier part of the decade in the American Community Survey (ACS) data. Both of these are similar in gains in the range of 20% (though a declining trend in 2000-2010).

The 1997 cohort of teachers have slightly lower abilities with the main significant drop in ability being around the 75th percentile of the conditional distribution, though the differences are relatively minor with insignificant mean differences. These findings however mask stark differences between demographic groups. White woman entering the teaching profession are of considerable lower ability in the 1997 cohort than the 1979 cohort. On the other hand minority women entering the teaching profession are of surprisingly higher ability in the 1997 cohort as compared to the 1979 cohort. When these effects are decomposed into cohort wide changes and relative changes it becomes evident increases in overall ability in the 1997 cohort have helped to mitigate the even larger relative drop in ability for white female teachers. Conversely the huge gains in ability for minority women is equally due to cohort wide increases and an increase in relative ability of teachers.

A bit of a puzzle emerges as there are strong negative trends in ability within the 1979 cohort, which, had it continued, should have led to large drops in ability of teachers between cohorts. The fact that this does not show up in the cohort comparison indicates there was a shift in trend in the later 1990's, though this data is not available. An explanation of this observation, which is beyond the scope of this paper, perhaps lies in economy wide shifts that occurred in the 1990s, such as the halting in the labor supply shift of women which was strong in the 1980s and documented by Blau and Kahn (2007).

References

[1] Altonji, Joseph, Prashant Bharadwaj and Fabian Lange (2008) "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," *NBER Working Paper*. No. 13883.

[2] Bacolod, Marigee (2007) "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality," *The Review of Economics and Statistics*. 89, 4, 737-751.

[3] Blau, Francine and Lawrence Kahn (2007) "Changes in the Labor Supply Behavior of Married Women: 1980-2000," *Journal of Labor Economics*. 25,393-438.

[4] Borman, Geoffrey and N. Maritza Dowling (2008) "Teacher Attrition and Retention: A Meta-Analytic and Narrative Review of the Research," *Review of Educational Research*. 78, 3, 367-409.

[5] Corcoran, Sean, William Evans and Robert Schwab (2004) "Women, the Labor Market, and the Declining Relative Quality of Teachers," *Journal of Policy Analysis and Managment*. 23, 3, 449-470.

[6] Corcoran, Sean, William Evans and Robert Schwab (2004) "Changing Labor Market Opportunities for Women and the Quality of Teachers," *American Economic Review: Papers and Proceedings*. 94, 2, 230-235.

[7] Flyer, Fredrick and Sherwin Rosen (1997) "The New Economics of Teachers and Education," *Journal of Labor Economics*. 15, 1, S104-S139.

[8] Hoxby, Minter and Andrew Leight (2004) "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States," *American Economic Review: Papers and Proceedings*. 94, 2, 236-240.

[9] Lakdawalla, Darius (2001) "The Declining Quality of Teachers," *NBER Working Paper*. No. 8263.

[10] Rice, Jennifer King (2003) "Teacher Quality: Understanding the Effectiveness of Teacher Attributes," Washington, DC: Economic Policy Institute.

[11] Segall, D. O. (1997) "Equating the CAT-ASVAB. In W. A. Sands, B. K. Waters, J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (pp. 181-198). Washington, DC: American Psychological Association.