**The Returns to Ability and Experience in High School Labor Markets: Revisiting Evidence on Employer Learning and Statistical Discrimination**

Xizi Li and Stephen L Ross, University of Connecticut

Abstract

We extend the empirical model of Arcidiacono, Bayer and Hizmo (2010) on employer learning and document a non-linear relationship between wages and ability (AFQT) at low experience levels. The return to AFQT is strong at low ability levels, but flat or decreasing at high levels. Much of the observed increase in return to AFQT as potential experience increases is associated with a change in the shape of this non-linear relationship. We also find that high AFQT workers without four years of college select into occupations that provide more training, perhaps sacrificing initial wages to build skills.

Key Words: Wages, Human Capital, Ability, Screening, Signaling, Learning, Statistical Discrimination, AFQT, Education, Compensating Differential, Training, Occupation, NLSY

JEL Codes: D82, D83, I24, I26, J24, J31

Acknowledgements: The authors are grateful to Joseph Altonji for providing the data used in A&P and to the authors of ABH for sharing their programs on-line. The manuscript benefited substantially from comments provided by Joseph Altonji, Peter Arcidiacono, David Deming, Caroline Hoxby, Dean Lillard, William Strange and the participants of the Spring 2019 Housing Urban Labor Macro meetings. All remaining errors are our own.

Corresponding Author: Stephen L Ross, 365 Fairfield Way U-1063, Storrs, CT 06269-1063, stephen.l.ross@uconn.edu, 860-486-3533.

**The Returns to Ability and Experience in High School Labor Markets: Revisiting Evidence on Employer Learning and Statistical Discrimination**

**I. Introduction**

The classic model of statistical discrimination implies that education serves as a signal for unobserved ability (Spence 1973; Weiss 1995). Altonji and Pierret (2001) provide evidence of statistical discrimination followed by employer learning. They use the National Longitudinal Survey of Youth (NLSY 79) to show that the return to ability as measured by the Armed Forces Qualification Test (AFQT) is low when workers first enter the labor market and as workers gain experience the correlation between wages and ability grows and the correlation between wages and education falls.[[1]](#footnote-1) Following up on their study, Arcidiacono, Bayer and Hizmo (2010), here after referred to as ABH, show that ability and initial wages are unrelated for high school graduates (exactly 12 years of education), but that ability appears to influence initial wages for college graduates (exactly 16 years of education).[[2]](#footnote-2) Similarly, Lange (2007), Kahn and Lange (2014), Light and McGee (2015a) and Dustmann, Glitz, Schonberg and Brucker (2016) provide evidence related to employers learning about worker ability. Further, MacLeod et al. (2017) show that in the country of Columbia implementation of an exit exam reduced the wage return to college reputation, presumably by providing information and reducing the value of the signal.[[3]](#footnote-3)

We use the analysis by ABH as a starting point given their discovery that effects are limited to those who do not have a four year college degree. We replicate the basic AFQT results in ABH for the high school degree sample and also show that both the AFQT and Altonji and Pierret’s (2001) education results arise for a sample of workers with less than exactly 16 years of education, i.e. four-year college degree. We then extend these models to include the square of AFQT in order to allow for a non-linear relationship between ability and wages. These models suggest that ability is related to initial wages, but in a non-linear manner with wages flat or declining with worker test score for above average test score workers. Next, we document how this non-linear relationship changes over time by estimating models of worker average wages for four year periods of potential experience. For between 1 and 4 years of potential experience average wages continue to exhibit a non-monotonic relationship with AFQT for above average AFQT scores. However, for later years of potential experience, the non-linear relationship begins to disappear, and a monotonic relationship between wages and AFQT develops. This non-linear relationship between AFQT and wages is robust to non-parametric modelling of the relationship, adding controls for sibling wage as developed in Altonji and Pierret (2001), allowing for non-linear returns to education, and finally is concentrated primarily among white workers when the sample is split by race following Pinkston (2006).

As with ABH’s results, the non-linear relationship between wages and AFQT is concentrated among our sample of workers who do not complete four years of college. This feature of our results raises concerns that high AFQT individuals who do not complete four years of college may represent a heavily selected sample. In order to test for this, we estimate a model of the likelihood of completing four years of college by the time of labor market entry, and then include a quadratic in the predicted likelihood of completion exploiting functional form to run a horse-race between the AFQT test score and the likelihood controls that provide a better proxy for the degree of selection in the sample. The estimates are consistent with the non-linearity being associated with cognitive ability rather than selection into the non-college sample.

Next, we use all available waves of the NLSY 79 to examine wages of a broader sample of workers without four years of college that encompasses both the Altonji and Pierret (2001) and ABH samples and uses the worker’s initial, pre-market level of education. We then repeat this exercise using the NLSY 1997. Again both using all waves of the NLSY 79 and using the NLSY 97, we continue to observe a non-linear relationship between AFQT and wages at low levels of potential experience that becomes monotonic, or at least closer to monotonic in the NLSY 97, as potential experience increases. In terms of returns to education, the evidence of declines in the return to education from years of potential experience is significantly weaker and less robust in the NLSY 79 after allowing for this non-linearity, while the NLSY 97 estimates suggest an increasing return to education with potential experience.

In principle, we cannot rule out a high school labor market where firms observe ability relatively accurately when workers have lower levels of ability, but cannot distinguish between workers with above average and high levels of ability and so statistically discrimination based on education when workers have above average levels of ability. This limitation is especially true for the results based on all waves of the NSLY 1979 in that the return to AFQT at high levels of AFQT is very close to zero at low levels of potential experience. However, the results based on the NLSY 97 work against this explanation because in that sample wages decline substantially with AFQT at high levels of AFQT when experience levels are low, and the wage return to education increases with potential experience.

In trying to explain these findings, we speculate that, just as high ability workers tend to attend college postponing current earnings for higher earnings later, perhaps high ability workers who do not obtain four years of college follow a similar pattern by taking jobs that pay less now, but contribute to the accumulation of human capital over time (Ben-Porath 1967). In fact, Altonji and Spletzer (1991) find a positive relationship between worker aptitude and training. To examine this premise in our sample, we estimate models related to the training received by workers early in their work career. Among workers who do not initially complete four years of college, high AFQT workers in the both the NLSY 79 and the NLSY 97 are more likely to receive firm provided training, and for the NLSY 79 are more likely to select initial occupations that tend to provide more firm training. Further, we re-estimate our potential experience subsample wage models including controls for exposure to occupations with high incidence of training, as a proxy for ability to accumulate skills. These controls explain much of the increasing return to AFQT for high AFQT workers as they gain potential experience. In the NLSY 79, the effect arises because high AFQT workers are more likely to be found in high training incidence occupations, while in AFQT 97 the effect arises because high AFQT worker appear to benefit more from exposure to high training incidence occupations.[[4]](#footnote-4)

As a whole, our findings appear consistent with firms being able to reward ability,[[5]](#footnote-5) but wages of high ability workers without out four years of college being depressed by the forward-looking choices of those workers. These findings are also relevant to earlier research related to on-the-job training. While many studies support Ben-Porath’s (1967) prediction that wages rise with training,[[6]](#footnote-6) evidence for his second prediction that initial wages are lower in jobs that provide training is virtually non-existent. Most studies never report results on initial wages. Barron, Black and Lowenstein (1989) find no evidence of lower initial wages arguing that the failure to find evidence of lower wages is likely due to the sorting of higher ability workers into jobs that provide more training.[[7]](#footnote-7) Veum (1999) only finds evidence of lower starting wages for one of the four types of training considered and that finding is for training that is not paid for by the firm. By focusing on how wages vary non-linearly over ability, we document lower initial wages for the types of workers who are likely to receive and/or benefit from training and occupations that provide training.

**II. Replication and Initial Evidence on Non-linearities**

The ABH samples include wage information from non-Hispanic male respondents ranging from 14 years old to 21 years old at the time of the first wave of the NSLY 79 through the survey wave conducted in 2004. For this sample of workers, they include one observation for every wave starting with the first wave in which a worker reports not being in school. ABH follows Lange (2007) by calculating potential experience based on the first year that the individual leaves school and participates in the labor market, and so workers in the labor market as of wave 1 are omitted if the worker did not report a graduation year and did not provide sufficient information on work history prior to wave 1 of the NLSY for the calculation of potential experience. ABH also restrict the sample to worker observations with 13 or less years of potential experience citing a non-linearity in the wage relationship for higher levels of potential experience. Finally, ABH examines subsamples with exactly 12 and 16 years of education.

Following ABH, but focusing on a reduced form model in potential experience rather than an IV model for actual experience, a simple OLS model of wages ($y\_{it}$) for worker *i* in survey year *t* can be written as

$y\_{it}=S\_{it}β\_{1}+\left(S\_{it}\*P\_{it}\right)β\_{2}+X\_{it}β\_{3}+δ\_{t}+ +ε\_{it}$ (1)

where $S\_{it}$ is the vector of skill variables composed of AFQT, possibly the square of AFQT, the race dummy variable, and in some models years of education; $P\_{it}$ is years of potential experience; $X\_{it}$ is the vector of additional contemporaneous controls; and $δ\_{t}$ represents the year fixed effects. The additional controls in $X\_{it}$ include a cubic in potential experience, whether the worker lives in an urban area, and whether the worker was employed full or part-time. Standard errors are clustered at the worker level.

Table 1 presents the reduced form replication of ABH with column 1 using the same subsample with exactly 12 years of education, and column 2 presenting results using a broader sample of anyone with less than four years of college (less than 16 years of education). The return to AFQT increases with potential experience from a rate of near zero to a return of approximately 12% for a one standard deviation in increase in AFQT over a decade consistent with employer learning. We find similar, but somewhat smaller effects, only 10%, in column 2 for the sample of workers with less than four-years of college. The estimates in Column 2 also illustrate a negative relationship between potential experience and the returns to years of education with the return to a year of education falling by about 2 percentage points over a decade.

Columns 3, 4, 5 and 6 repeat these analyses just adding the square of AFQT as a regressor in the case of columns 3 and 4 and also adding the interaction of potential experience and the square of AFQT in columns 5 and 6. We find a statistically significant negative coefficient on the square of AFQT in all four columns suggestive of falling wages with AFQT for workers who have high levels of AFQT. The coefficient on the interaction of the square of AFQT with potential experience is very small and insignificant. The near zero estimate on the linear AFQT term at one year of potential experience in columns 3-6 suggests that initially the wage return to AFQT is negative for workers with an above average AFQT score. Further, even with an insignificant estimate on the interaction of potential experience and the square of AFQT, the non-linear relationship between AFQT and wages may change with potential experience because the estimate on the linear AFQT term increases with experience moving the value of AFQT where expected wages reach their to the right as workers gain experience.[[8]](#footnote-8) [[9]](#footnote-9)

In order to examine this more carefully, we divide the sample into observations with a similar number of years of potential experience: 1-4, 5-8, 9-12, 13-16; and then to reduce noise and measurement error in wages we collapse the data to the worker level in order to measure average wages at different levels of potential experience.[[10]](#footnote-10) A four-year window was selected in order to provide at least the potential for two years of wage data, i.e. years in which the individual was surveyed and working even after the NLSY changed from annual surveys to surveys every two years. This average wage is then regressed upon the controls for AFQT, square of AFQT, race and the within worker average for the subsample of the time varying controls in the ABH model including years of education as we focus on the sample of all workers with less than 16 years of education going forward. A separate model is estimated for each potential experience subsample and so robust standard errors are reported because each subsample has only one observation per worker.

Table 2 presents these results. We continue to find strong evidence of a non-monotonic relationship between wages and AFQT when workers have little potential experience. As potential experience increases, the coefficient on the linear term for AFQT increases in magnitude leading to a more monotonic relationship between wages and AFQT. The row labelled “turning point” shows the AFQT score where the slope of the relationship with wages is zero, and this turning point increases with potential experience. The slope changes are illustrated by Figure 1 Panel 1, which plots the estimated quadratic relationship between wages and AFQT for the model. The figure shows a humped shape relationship between wages and AFQT for low levels of potential experience that slowly approaches a more linear relationship as potential experience increases. While the return to AFQT at low levels of AFQT does appear to increase somewhat with potential experience, much of the increase in return to AFQT over the entire range of AFQT arises from the change in the shape of the curve, rather than an increase in the steepness over that range. Finally, the estimated coefficients on years of education fall by 1 percentage point at 9-12 years, but by less than ½ a percentage point at 13-16 years of potential experience, and the evidence of declining returns to education with low levels of potential experience is weaker.[[11]](#footnote-11)

Next, we examine non-parametric models of the relationship between log wages ($y\_{i}$) and AFQT score ($A\_{i}$) for each of our potential experience subsamples.

$y\_{i}=X\_{i}β+f(A\_{i})+ ε\_{it}$ (A1)

We first follow Robinson and obtain consistent estimates of $β$ by conditioning both $y\_{i}$ and the elements of the vector $X\_{i}$ on the non-parametric function of $A\_{i}$.

 $y\_{i}=g\left(A\_{i}\right)+μ\_{i}$ (A2)

 $X\_{i}=h\left(A\_{i}\right)+π\_{i}$ (A3)

and conducting a regression using the predicted residuals

$\hat{μ}\_{i}=\hat{π}\_{i}β+\tilde{ε}\_{i}$ (A4)

Then, we use the consistent estimates of $β$ to estimate

$y\_{i}-X\_{i}\hat{β}=f(A\_{i})+ \tilde{\tilde{ε}}\_{it}$ (A5)

We use non-parametric kernel regression techniques to both condition $y\_{i}$ and $X\_{i}$ on a non-parametric function of $A\_{i}$ and to estimate the non-parametric relationship between the wage residual and $A\_{i}$. Specifically, we calculate a weighted average of the left-hand side variable over the entire sample based on the distance in $A\_{i}$ space between the point of interest and each point in the sample. For $y\_{i}$ and $X\_{i}$, we conduct this exercise for every observation *i* in order to calculate the predicted residuals. For example,

$\hat{μ}\_{i}=y\_{i}-\sum\_{∀j}^{}y\_{j}α\left(\left|A\_{i}-A\_{j}\right|,ω\right)$ (A6)

where $α$ is a set of weights based on a triangular kernel with bandwidth $ω$ so that

 $α\left(\left|A\_{i}-A\_{j}\right|,ω\right)=\frac{1}{\sum\_{∀k}^{}\left(1-\frac{1}{ω}\left|A\_{i}-A\_{k}\right|\right)}\left(1-\frac{1}{ω}\left|A\_{i}-A\_{j}\right|\right)$ (A7)

For estimating the surface, we conduct the same exercise over a grid of values of AFQT indexed by $a$ in order to estimate the non-parametric relationship as

$f\left(A\_{a}|ω\right)=\sum\_{∀j}^{}\left(y\_{j}-X\_{j}\hat{β}\right)α\left(\left|A\_{a}-A\_{j}\right|,ω\right)$ (A8)

The bandwidth $ω$ is selected using a cross-validation approach. In the first stage for estimating equations (A2) and (A3), we select the bandwidth to minimize the mean squared error $\hat{μ}\_{i}$ in equation (A6) with the restriction that the summation in (A6) omits observation *i* for the cross-validation exercise. In the second stage, the bandwidth is selected to minimize the deviation of the wage residual from the estimated surface or equivalently to minimize the mean squared error of

$\hat{ε}\_{i}=\left(y\_{i}-X\_{i}\hat{β}\right)-\sum\_{∀j\ne i}^{}\left(y\_{j}-X\_{j}\hat{β}\right)α\left(\left|A\_{i}-A\_{j}\right|,ω\right)$ (A9)

For all waves of the NLSY 79, we find optimal bandwidths ranging between 0.5 and 1.1 across the potential experience subsamples with somewhat higher bandwidths in the second stage. The estimated bandwidths are shown Appendix Table A5.

 In order to practically estimate this semi-parametric model, we select a more somewhat more parsimonious model specification replacing the year fixed effects with a cubic polynomial in year. The estimated non-parametric relationships are presented in Figure 1 Panel 2. Again, for low levels of potential experience, the model illustrates a decline in wages with AFQT at the highest AFQT levels, which reverses and becomes positive as potential experience increases. The return to AFQT again also increases with potential experience at low levels AFQT, but a substantial portion of the increase in return to AFQT continues to arise from a change in the shape of the AFQT wage curve.

Some economists have informally raised concerns about the years of education sample restrictions in ABH. Given the concentration of our non-linearities in the less than 16 years of education subsample, one might be concerned about selection into the sample of workers with less than 16 years of college. Perhaps, the non-linear relationship arises because high AFQT individuals who do not complete four years of college are negatively selected on other factors reducing their wages, and the highest AFQT workers are most selected and so have the lowest wages. To test for this, we estimate a probit model for having an initial education of at least four years of college in the entire NLSY sample of workers that forms the basis of Table 3. As before, we use the same model controls as used in Table 4 except years of education is omitted and the year fixed effects and trends are replaced by fixed effects associated with the age of the individual at wave 1 since both years of education and year entering the labor market are endogenous to whether the individual has an initial education of at least four years of college.

We then use the model to predict the likelihood of having four years of college when entering the labor market, and include both this predicted probability and the square of the predicted probability in the wage models. This exercise exploits the functional form in the probability model in order to run a horse race between our measure of cognitive ability and this proxy for the degree of selection faced by each observation in our less than 16 years of education sample. These results are shown in Table 3. Focusing on column 1 for 1-4 years of potential experience. The non-linearity in AFQT is even stronger than in Table 2 with wages beginning to decline with AFQT at a standardized AFQT score of -0.20, as opposed to a turning point of 0.75 in Table 2.[[12]](#footnote-12)

We then conduct a series of robustness tests that are detailed in the Appendix. Altonji and Pierret (2001) examine an alternative measure of ability, sibling wage, and find similar relationships between wage returns and potential experience. Our non-linear relationship between the wages and AFQT is robust to the inclusion of controls for sibling wage. In fact, the estimate on sibling wage is very stable as the AFQT controls are added, and vice-versa. So, these results might be interpreted as finding of a robust non-linearity in the wage return to cognitive skills after controlling for a relatively independent measure of skills (Appendix Table A6).[[13]](#footnote-13) Following Pinkston (2006), we also estimate models separately by race. The non-linear relationship between wages and AFQT is again very robust for the white subsample. The estimates on the non-linear term for 1-4 years of potential experience are smaller and insignificant in the black sample (Appendix Table A7).[[14]](#footnote-14) Further, the concentration of evidence for statistical discrimination in the black subsample is also consistent with Lang and Manove’s (2011) finding that blacks are over-represented in higher education conditional on ability, and ABH’s argument that a potential explanation for the selection of blacks into higher education is their desire to avoid discrimination by revealing ability through success in college.[[15]](#footnote-15)

Finally, we conduct some additional analyses to examine the sensitivity of our results to how we condition on education. First, we estimate a model with the full ABH sample unrestricted on years of education allowing for the non-linear relationship for AFQT to differ by whether individuals have completed four years of college or not (Appendix Table A8).[[16]](#footnote-16) Again, the non-monotonic relationship between AFQT and wages for low levels of potential experience is robust for the subsample without four year of college. Finally, the findings are robust to allowing the return to years of education to be non-linear (Appendix Table A9).

**III. Expanding the NSLY Sample**

We now use all available waves of the NLSY 79 to generate a broader sample of workers and worker wages. We also expand the sample by relaxing some of the restrictions placed upon the sample by ABH. ABH dropped workers who were not in school in wave 1 if the survey responses did not contain sufficient information to determine when they entered the labor market. We retain those individuals and calculate potential experience for those workers using the standard measure of age minus years of education minus six. ABH also deletes observations when the wage in the most recent job is invalid. However, each wave from the NLSY can in principle collect wages from multiple jobs, and we follow Altonji and Pierret (2001) use the first reported valid wage in the survey if the wage in the most recent job is unobserved. Our sample of workers includes all non-Hispanic white and black, male workers who have at least 8 years of education and a valid wage in at least one wave when not in school, and then we include observations for all future waves in which they have valid wages. Our final sample includes virtually every worker included in either ABH or Altonji and Pierret (2001), as well as an additional 700 workers who were in neither sample.

Next, in a key departure from ABH, we focus on the education level of the individual when they either first left school or on their wave 1 education level if they were not in school at wave 1 (initial education) in order to have a more exogenous measure of education. The sample is restricted to all workers who had less than 16 years of education initially upon leaving school, and the control for current years of education is replaced by a control for initial years. Then, potential experience for most of the sample is simply the number of years since they first left school, but in some cases potential experience is age minus initial years of education minus six. In both cases, potential experience is unaffected by any events or decisions that arise after initial education is observed.[[17]](#footnote-17) Finally, we observe that many individuals in school at wave 1 appear to leave school or graduate in a later wave (report not being in school), and yet continue to increment their years of education by exactly one year in every following year for one or more years after the reported date of leaving school. In those cases, we set the first year of potential experience equal to the first year that years of education does not increase, and use the years of education observed in that year as the initial years of education.[[18]](#footnote-18) Finally, we continue to allow wages to depend upon whether the worker is currently employed part-time and whether the worker resides in an urban area.

We then re-estimate wage models from Table 2 using this larger sample of workers, the initial years of education and potential experience based on that initial level and again collapsing the data into subsamples of workers based on wage observations for similar years of potential experience. Table 4 Panel 1 presents the estimates for the full sample of individuals without 16 years of education, and Panel 2 presents the same model estimates for the white subsample given the differences between whites and blacks described above. Each column presents estimates for average wages for four specific years of potential experience up to 29 to 32 years. Figure 2 Panels 1 and 2 present plots of the estimated quadratic relationship between wage and AFQT for the full and white only sample, and Panel 3 presents the non-parametric relationships for the full sample. The non-monotonic relationships in Figure 2 for low levels of potential experience are robust and very similar to the plots in Figure 1. In all panels, a substantial portion of the increasing return to AFQT with potential experience is attributable to the changes from a non-monotonic to a monotonic relationship between AFQT and wages. In fact, for the white sub-sample (panel 2) and the non-parametric models (panel 3),[[19]](#footnote-19) the slope of the AFQT-wage curve for 1-4 years of potential experience is virtually indistinguishable from the curves for higher levels of potential experience at low levels of AFQT. Further, estimates on the return to years of education are relatively stable across the subsamples, although somewhat noisier in the white only samples.

Next, we create a comparable sample using workers for the NLSY 97 imposing the same sample restrictions described above. Table 5 presents estimates of the return to AFQT for workers with less than four years of college and 1-4, 5-8, 9-12 and 13-16 years of potential experience using data from the NLSY 97 and the same model controls as used in Table 4.[[20]](#footnote-20) The estimated AFQT coefficients are also illustrated in Figure 3 for the full sample, the white worker only sample and for the full sample using the non-parametric analysis.[[21]](#footnote-21) The non-linearity in return to AFQT appears stronger in the NLSY 97. For low levels of potential experience between 1 and 8 years, the relationship between AFQT and wages is clearly negative for the entire upper half of the AFQT distribution. The figure suggests a substantial strengthening of the return to AFQT for low levels of AFQT as potential experience increases, but even at potential experience between 13 and 16 years wages do not appear to increase in AFQT for high levels of AFQT. Further, returns to years of education increase with potential experience in the NLSY 97.[[22]](#footnote-22) [[23]](#footnote-23)

In summary, neither broadening the sample, adding information from the additional waves of data, using data from younger cohorts at a later time changes our basic conclusions. First, the return to AFQT is non-monotonic for low levels of potential experience for workers without four years of college and much of the increase in the average return to potential experience can be attributed to the shift towards a more linear relationship as potential experience increases. Second, declines in the return to education with potential experience tend to be less robust than those identified by Altonji and Pierret (2001).

**IV. Are High AFQT Workers Building Skills?**

It is well accepted that high ability workers tend to attend college, postponing earnings early in life in exchange for building human capital, and so receiving higher earnings later in life.[[24]](#footnote-24) Perhaps, high ability workers who do not go to college follow a similar pattern by taking jobs that pay less now, but contribute to the accumulation of human capital over time and so earn higher wages later in life (Ben-Porath 1967). To examine this possibility, we look at firm provided training. Kahn and Lange (2014) show that a substantial portion of wage dispersion is attributable to worker learning and associated productivity gains. If learning is facilitated by cognitive ability, high AFQT workers may prefer jobs that convey substantial skills over time, especially if they have not invested in such skills through higher education (Barron, Black and Lowenstein 1989).

We develop a measure of training as the fraction of years in the labor market in which an individual reports receiving firm provided training during the first four years after entering the labor market. We regress this measure for our sample of workers with less than 16 years of initial education controlling for the standard controls from our 1-4 years of potential experience subsample wage model specification in Table 4. These results are shown in columns 1 and 2 of Table 7 with column 1 presenting results for firm sponsored training and column 2 presenting results for firm sponsored training plus apprenticeships. Panels 1 and 2 present results for our NLSY 79 expanded sample of workers and our equivalent sample of workers from the NLSY 97, respectively. A one standard deviation change in AFQT is associated with an increase in the share of years receiving training of between 1/3 and 1/2 of the average training share experienced in the sample. These results are similar to those of Altonji and Spletzer (1991) that workers with higher Math SAT scores experience more training.

Next, we use the entire sample of either NLSY 1979 or NLSY 1997 workers to identify the occupations where firm sponsored training or apprenticeships are most common. Specifically, we calculate the share of years receiving training during the first four years of work for the entire NLSY sample of male workers, relaxing many of the sample restrictions placed on our regression sample, and then calculate the average of the training by occupation. We assign workers in our regression sample to a training intensity based on their initial job’s occupation omitting themselves from the calculation of intensity for their occupation, and then as an alternative we also assign workers based on their occupation at five years (or six years if occupation unobserved at five years) in order to address the possibility that it takes time for workers to sort into their optimal occupation. Columns 3 and 4 present results for initial occupation, and columns 5 and 6 use occupation at five years. In the NLSY 1979, Higher AFQT workers appear to select into occupations that provide more training. A one standard deviation increase in AFQT is associated with an initial occupation choice into occupations where average training share is 10 percent higher than the training share observed in the average occupation, and this effect increases to about 20 percent for occupation at five years. The effect of initial occupation in the NLSY 97 is less 1/2 the magnitude in the NLSY 79, and the effect disappears completely for occupation at five years of potential experience.[[25]](#footnote-25)

Next, we treat being in an occupation with a high incidence of employer providing training as a proxy for opportunities for workers to build skills during employment, recognizing that the self-reported incidence of firm provided training is too small to explain a significant portion of wage variation. Specifically, we re-estimate the wage models in Tables 4 and 5 for the NLSY 1979 and 1997 adding controls for cumulative exposure at the time wages are observed to occupational training intensity, the square of cumulative exposure to allow for diminishing returns and the interaction of cumulative exposure with AFQT to allow for the possibility that high AFQT workers benefit more from opportunities to build skills during employment. Cumulative exposure calculated by simply adding up the incidence of training in a worker’s occupation each year over all years worked prior to a given wage observation.[[26]](#footnote-26)

These results are shown in Table 7 for the NLSY 1979 and in Table 8 for the NLSY 1997. The first panel repeats the estimates from Tables 4 and 5 for comparison purposes, and the second panel presents the estimates for the cumulative exposure model. In both samples, cumulative exposure to occupational training intensity is associated with higher wages, but also exhibits diminishing returns in cumulative exposure. In the NLSY 1979, the return to cumulative exposure does not depend upon AFQT. However, since high AFQT workers are concentrated in high training exposure occupations the wage return to AFQT in panel 2 does not increase with potential experience and cumulative exposure appear to explain most of the changes in return to AFQT as workers gain experience. In the NLSY 97, the return to cumulative exposure increases with AFQT, and the higher rate of return implies again that cumulative exposure can explain most of the changes in return to AFQT even though high AFQT workers are not disproportionately represented in those high training intensity occupations.[[27]](#footnote-27)

Finally, we investigate two other potential mechanisms for the non-linear relationship between initial wages and AFQT for workers who do not complete four years of college. Perhaps high AFQT workers are directly sorting into initial occupations with high levels of wage growth or sorting into higher wage occupations as potential experience increases. For this analysis, we only focus on the NLSY 1979 because we did not observe sorting across occupations in the NLSY 1997 in Table 6. We use the Current Population Survey to characterize occupation wage growth and levels over potential experience, and estimate models in the NLSY where these growth or level estimates are the dependent variables. However, as shown in Appendix Table 16, we do not find any relationship between initial or early occupation growth rates and AFQT. We do find that high AFQT workers move into higher wage occupations as they gain potential experience in the full sample, but these results are not robust to restricting our sample to white workers only, the subsample where our non-linear returns are concentrated, see Appendix Table 17. Therefore, we cannot replicate the effects of occupational training levels by simply examining occupation specific wage levels and/or wage growth rates.

**V. Discussion**

 In summary, the signaling model of educational investment is a very important theory within labor economics. Much of the evidence in support of this theory is based on analyses of the 1979 National Longitudinal Survey of Youth where initial wages have been found to be unrelated to a test score proxy for ability, but to increase with test score based on time in the labor market, while simultaneously the importance of education in explaining wages falls with time. In this paper, we show that AFQT is related to early wages for workers who did not complete four years of college, but in a non-linear manner with wages rising with AFQT for workers with low scores and falling with AFQT for workers with high scores. This non-linear relationship disappears as workers obtain more experience. These patterns are very robust across samples and model specifications, and a substantial portion of the increase in the return to AFQT with potential experience in the NLSY is likely due to the short-run nature of this non-monotonic relationship. Further, as previously observed by Pinkston (2006), almost all of the evidence that we identify for declining returns to education with potential experience is not present in a model that excludes black workers.

The evidence appears consistent with the highest ability workers who do not complete four years of college investing in higher future earnings by selecting jobs that build human capital. Workers without four years of college, but with high AFQT, are observed to receive more training than low AFQT workers and to select into occupations where training is provided more frequently. Further, exposure to high training intensity occupations appears to explain most or all of the increase in the return to AFQT that arises among workers without four years of college. It is important to note, however, that when just focusing on low AFQT workers the return of AFQT in wages does increase with potential experience in some models. Therefore, statistical discrimination may still play a role in explaining the early wages of workers, but past studies that have not considered compensating wage differentials based on future skill accumulation may significantly overstate the importance of signaling for demonstrating cognitive skills in the labor market.

**Reference**

Acemoglu, Daron and Jorn-Steffen Pischke. (1999). The Structure of Wages and Investment in General Training. *Journal of Political Economy*, 107(3), 539–572.

Altonji, Joseph G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics,* 116(1):331-350.

Altonji, Joseph G. and James R. Spletzer. (1991). Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training. *Industrial and Labor Relations Review*, 45(1), 58-79.

Arcidiacono, Peter, Bayer, Patrick, & Hizmo, Aurel. (2010). Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics*, *2*(4), 76-104.

Arteaga, Carolina. 2018. The effect of human capital on earnings: Evidence from a reform at Colombia's top university. *Journal of Public Economics*, 157, 212-225

Bacolod, M., & B. S. Blum (2010). Two Sides of the Same Coin: U.S. “Residual" Inequality and the Gender Gap. *Journal of Human Resources*, 45(1), 197–242.

Barron, John M., Dan A. Black and Mark A. Loewenstein. (1989). Job Matching and On-the-Job Training. *Journal of Labor Economics*, 7(1), 1-19.

Bartel, Ann P. (1995) Training, Wage Growth, and Job Performance: Evidence from a Company Database. *Journal of Labor Economics* 13(3), 401-425.

Bartel, Ann P., & Sicherman, N. (1998). Technological change and the skill acquisition of young workers. *Journal of Labor Economics*, 16(4), 718-755.

Ben-Porath, Yoram. 1967. The Production of Human Capital and the Life- Cycle of Earnings. *Journal of Political Economy*, 75, 352-65.

Bjerk, David. (2007). The Differing Nature of Black-White Wage Inequality Across Occupational Sectors. *Journal of Human Resources*, 42(2), 398-434.

Booth, Alison L. and Mark L. Bryan. (2005). Testing Some Predictions of Human Capital Theory: New Training Evidence from Britain. *Review of Economics and Statistics*, 87, 391–394.

Brown, James N. (1989). Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms. *American Economic Review*, 79, 971-91.

Castex, Gonzalo and Evgenia Kogan Dechter. (2014). The changing roles of education and ability in wage determination. *Journal of Labor Economics*, 32(4), 685-710.

Dearden, Lorraine, Howard Reed, John Van Reenen. (2006). The Impact of Training on Productivity and Wages: Evidence from British Panel Data. *Oxford Bulletin of Economics and Statistics*, 68(4), 397-421.

Deming, David. and Kadeem Noray. (2018). STEM Careers and Technological Change. NBER Working Paper #25065.

Dustmann, C., Glitz, A., Schönberg, U., & Brücker, H. (2015). Referral-based job search networks. *The Review of Economic Studies*, *83*(2), 514-546.

Fang, Hanming. (2006). Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review*, *47*(4), 1151-1185.

Farber, H. S., & Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, *111*(4), 1007-1047.

Kahn, Lisa B. and Fabian Lange. (2014). Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures. *Review of Economic Studies*, 81(4), 1575-1613.

Konings, Jozef and Stijn Vanormelingen. 2015. *The Review of Economics and Statistics*, 97(2), 485–497.

Lang, Kevin and Michael Manove. (2011). Education and Labor Market Discrimination. *American Economics Review* 101, 1467-1496.

Lange, Fabian. (2007). The speed of employer learning. *Journal of Labor Economics*, *25*(1), 1-35.

Light, Audrey and Andrew McGee. (2015a). Employer Learning and the “Importance” of Skills. *Journal of Human Resources*, 50(2), 72-107/

Light, Audrey and Andrew McGee. (2015b). Does employer learning vary by schooling attainment? The answer depends on how career start dates are defined. *Labour Economics*, 32(1), 57-66.

Loewenstein, Mark A. and James R. Spletzer. (1998). Dividing the Costs and Returns to General Training. *Journal of Labor Economics*, 16(1), 142–171.

Lynch, Lisa M. (1992). Private-Sector Training and the Earnings of Young Workers. *American Economic Review*, 82, 299–312.

MacLeod, W. Bentley, Evan Riehl, Juan E. Saavedra and Miguel Urquiola. (2017). The Big Sort: College Reputation and Labor Market Outcomes. *American Economic Journal: Applied Economics*. 9(3), 223-261.

Pinkston, Joshua. (2006). A test of screening discrimination with employer learning. *Industrial and Labor Relations Review*, 59(2), 267-284.

Spence, Michael. (1978). Job market signaling. In *Uncertainty in Economics* (pp. 281-306).

Veum, Jonathan. (1999). Training, Wages, and the Human Capital Model. *Southern Economic Journal*, 65(3), 526-538.

Weiss, A. (1995). Human capital vs. signaling explanations of wages. *Journal of Economic perspectives*, *9*(4), 133-154.

**Figure 1. Estimated Relationship between AFQT and Average Wages ABH Sample**





**Figure 2. Relationship between AFQT and Average Wages Expanded NLYS 1979 Sample**

****

****

****

**Figure 3. Relationship between AFQT and Average Wages NLSY 1997 Sample**

****





1. Also see earlier work on this topic by Farber and Gibbons (1996). [↑](#footnote-ref-1)
2. In fact, our own analysis shows that the once the additional waves in ABH are added the effects identified by A&P cannot be detected statistically in the full sample likely due to the larger number of college educated workers in the sample. [↑](#footnote-ref-2)
3. Arteaga (2018) shows the opposite effect that workers are paid for skills at hiring. Specifically, they document a decline in the wages of graduates when course work requirements for the degree fell. Her findings could also be consistent with ABH’s finding that college reveals ability through grades and the coursework completed. [↑](#footnote-ref-3)
4. On the other hand, we do not find robust evidence that these workers simply sort into initial occupations with higher wage growth or sort into occupations with higher wages as experience increases. [↑](#footnote-ref-4)
5. Consistent with Fang’s (2006) finding that most of the return to education is due to the value of skills, as opposed to signaling. [↑](#footnote-ref-5)
6. See for example Bartell (1995), Booth and Bryan (2005), Brown (1989), Dearden, Reed and Reenan (2006), Konings and Vanormelingen (2015), Lowenstein and Spletzer (1999) and Lynch (1992). [↑](#footnote-ref-6)
7. Acemoglu and Pischke (1999) offer an alternative explanation that firms finance general skills because labor market frictions convert general skills to specific. Considerable empirical work suggests that firms capture some of the return to providing general skills, but virtually all of that work also has a considerable share of returns accruing to workers suggesting that workers should be willing to forego initial earnings to obtain those skills (Booth and Bryan 2005; Konings and Vanormelingen 2015; Loewenstein and Spletzer 1999). [↑](#footnote-ref-7)
8. Altonji and Pierret (2001), here after referred to as A&P, estimate a model that also includes occupation fixed effects and allows the return to education and AFQT and the effect of race to vary based on the year that the individual is observed in the labor market. Appendix Tables A1 and A2 present a replication of A&P and replicates ABH using models that include controls for occupation and/or trends in the wage returns, respectively. We continue to find robust evidence of a non-linear return to AFQT at low levels of potential experience in all models estimated. [↑](#footnote-ref-8)
9. As in ABH, we do not find any evidence of increasing return to AFQT or decreasing return to education with potential experience in samples of workers with either 16 years of education or 16 or more years of education. Further, we do not find any evidence of non-linear returns to AFQT in these four years of college samples. The pooled sample tends to provide estimates that are between the estimates arising from the two subsamples [↑](#footnote-ref-9)
10. Note that we add a subsample of workers with 13-16 years of potential experience, which were excluded from ABH’s original sample, since we are estimating separate regressions for each subsample. [↑](#footnote-ref-10)
11. Appendix Table A3 and Figure A1 present similar analyses using the A&P sample. Also, Appendix Table A4 presents estimates for the sample with 16 or more years of education. We do not observe the non-linear relationship between wages and AFQT, and we observe an increasing return to years of education with potential experience for this sample. [↑](#footnote-ref-11)
12. Similar results arise if we simply split the analysis by terciles of the likelihood of having four years of college or more. In that case, we observe that the non-linearity for workers with less than four years is largest in the bottom tercile where 98 percent of the workers do not have four years of college. While the tercile analysis might appear as if it relies less on functional form, in fact the bottom tercile has minimal support over the region where wages fall with AFQT. Therefore, we believe that the horse race above is a more reasonable approach because the models examines the AFQT relationship by comparing observations with similar levels of selection over the entire support of the sample. [↑](#footnote-ref-12)
13. While the wage return to sibling wage increases in importance as we move from the 1-4 years of potential experience subsample to the 5-9 years subsample. We do not observe a decline in the return to years of education as potential experience increases, inconsistent with the signaling through education for the skills measured by sibling wage. Finally, we do not present models including the square of sibling wage because estimates on the square are always insignificant and unlike with AFQT including the square term also leads to very noisy estimates on sibling wage effects in general. [↑](#footnote-ref-13)
14. Again, we find little or no evidence of a decline in the return to years of education for the white subsample. These results are consistent with Pinkston’s (2006) finding that statistical discrimination on education was primarily experienced by black workers. [↑](#footnote-ref-14)
15. Similarly, Bjerk (2007) finds that ability explains the entire black-white wage gap for workers in white collar jobs, while wage gaps persist in blue collar jobs, and that conditional on ability blacks are more likely to work in white collar jobs. [↑](#footnote-ref-15)
16. If one is willing to condition on years of education as a right hand side variable, then in principle one should be willing to condition the sample on years of education as well. So, perhaps some of the concern with restricting the sample based on years of education arises because employers may be using completion of specific education thresholds as a signal and the effect of these thresholds on wages is lost when the sample is restricted. Therefore, the model also includes controls for years of education, completion of at least 12 years of education and completion of at least 16 years of education plus the standard controls in the ABH model. The total returns to education appear relatively stable over years of potential experience. [↑](#footnote-ref-16)
17. ABH describe in their paper not incrementing potential experience when an individual leaves the labor market and returns to school, but in reviewing their code their potential experience also appears to increment every year regardless of whether the worker returns to school. [↑](#footnote-ref-17)
18. Light and McGee (2015b) also document this issue and address it in their models. [↑](#footnote-ref-18)
19. For the subsamples using all waves of the NLSY 79, we find optimal bandwidths ranging between 0.7 and 1.1 across the potential experience subsamples with somewhat higher bandwidths in the second stage, again with higher bandwidths in the second stage, see Appendix Table 5. [↑](#footnote-ref-19)
20. Appendix Table 10 presents interactive models similar to those of ABH and A&P for the NLSY 97. The non-linear relationship between wages and AFQT arises for workers without four years of college in those models as well. As in ABH, the estimates imply increasing returns to AFQT as potential experience increases, but we do not observe declining returns to years of education as potential experience increases with this model. [↑](#footnote-ref-20)
21. For the NLSY 97, we find optimal bandwidths between 0.9 and 2.0, again with higher bandwidths in the second stage, see Appendix Table A5. [↑](#footnote-ref-21)
22. Like Castex and Dechter (2014), some basic patterns in wages documented in the NSLY 79, specifically those consistent with statistical discrimination followed by employer learning, do not persist in the NSLY 97 for either our full sample or the sample with less than four years of college. [↑](#footnote-ref-22)
23. Appendix Tables 11 and 12 presents results allowing for A&P style cubic year trends on the wage returns to AFQT, education and race for the the NLSY 79 and 97 samples, respectively. [↑](#footnote-ref-23)
24. See Fang (2006) for example. [↑](#footnote-ref-24)
25. Results are similar for the white worker only subsample. See Appendix Table 13. [↑](#footnote-ref-25)
26. As with time varying urban and part-time variables, this variable is an average over all wage observations included in the given potential experience subsample for each worker. [↑](#footnote-ref-26)
27. Results are similar for the white worker only sample. See Appendix Tables 14 and 15. [↑](#footnote-ref-27)