Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

# Discussion Papers

1015

Hani Mansour

# Does Employer Learning Vary by Occupation?

Berlin, June 2010

IMPRESSUM



DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN.
Papers can be downloaded free of charge from the following websites:

**http://www.diw.de/de/diw_01.c.100406.de/publikationen_veranstaltungen/publikationen/diskussionspapiere/diskussionspapiere.html**
**http://ideas.repec.org/s/diw/diwwpp.html**
**http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991**

# Does Employer Learning Vary by Occupation?

Hani Mansour

*University of Colorado Denver and DIW Berlin**

May, 2010

**Abstract**

Models in which employers learn about the productivity of young workers, such as Altonji and Pierret (2001), have two principal implications: First, the distribution of wages becomes more dispersed as a cohort of workers gains experience; second, the coefficient on a variable that employers initially do not observe, such as the Armed Forces Qualification Test (AFQT) score, grows with experience. If employers' learning varies significantly across occupations, both of these indicators of learning should covary positively across groups defined by a worker's occupational assignment at labor market entry. This paper tests this implication of the employer learning model using data from the NLSY and CPS. I find that occupations with high growth in the variance of residual wages over the first ten years of the worker's career are also the occupations with high growth in the AFQT coefficient, confirming the learning perspective. Interestingly, occupations that my analysis characterizes as having a low level of employer learning are not occupations where employers know little about the worker after ten years of experience; instead they appear to be occupations where employers have already learned about the worker's AFQT score at the time of hire. I provide several pieces of evidence that occupational assignment affects the learning process independently from education and that the results are not driven by workers' occupational mobility.

JEL Codes: J24, J31, J71, J62

Keywords: Wage Dynamics, Occupational Choice, Earnings Inequality.

---

*I am grateful to Peter Kuhn for his helpful comments and continuous support. I also thank Kelly Bedard, Olivier Deschênes, Peter Rupert, Javier Birchenall, Catherine Weinberger, David Card, and labor lunch participants at the University of California, Berkeley for their valuable comments and suggestions.

# 1 Introduction

Employers often hire new inexperienced workers without observing their full productivity. Instead, they assess a worker's value to the firm based on information they receive from job interviews, resumes, and recommendations (Spence, 1973). Altonji and Pierret (2001) [hereinafter AP] show that if employers learn about the worker's productivity, the coefficient of an ability correlate which is initially unobserved by employers, such as the Armed Forces Qualification Test (AFQT) score, increases with experience. AP's analysis, and most other models that use a learning framework, assume that learning is independent of job assignment. It is likely, however, that employers' learning varies significantly across occupations.[1]

Another principal implication of employer learning is that the distribution of wages becomes more dispersed as a cohort of workers gains experience (Neal and Rosen, 1998). If in fact employers' learning varies across occupations, both of these indicators - the growth in the AFQT coefficient and the growth in wage dispersion should covary positively across groups defined by a worker's occupational assignment. This paper tests this implication of the employer learning model and provides evidence confirming the learning perspective.[2]

Conceptually, learning might differ across occupations because of *level* differences in the variance of individual ability generated from non-random sorting into occupations or because the technology or tasks in each occupation affect the *speed* at which employers learn.[3] The

[1]An exception is Altonji (2005) who presents a framework in which the rate of employer learning depends on the skill level of the job to show how, in this environment, statistical discrimination at the time of hire affects employment rates and wage growth, but he does not test the model empirically. In addition, Antonovics and Golan (2007), motivate the idea of "job shopping" within firms using a model in which learning differs across jobs.

[2]Matching models, such as Jovanovic's (1979a) model, also generate increased wage dispersion with experience. Miller (1984) discusses an environment where workers learn about the quality of their match at different rates across occupations. The analysis I use help in distinguishing between the learning and occupational matching hypotheses.

[3]Consider the following example: Suppose that the true cross-sectional productivity variance in occu-

main contribution of the paper is to provide evidence that initial occupational assignments are associated with different learning parameters. The occupational analysis provides several pieces of evidence that distinguish the learning hypothesis from other competing hypotheses, such as on-the-job training (OJT), and improved match quality with experience. The results have important implications for various models in labor economics that use such frameworks. These include models of statistical discrimination, earnings inequality, occupational mobility, wage dynamics within firms, occupational wage differences, and labor market signaling.[4]

The model in the paper follows closely that of AP and Lange (2007) [hereinafter Lange].[5] Identical employers form expectations about the worker's productivity, and in each period update their initial belief based on a noisy signal of output produced by the worker. Lange shows that the speed of employer learning depends on the variance of the initial error that employers have and the variance of the noisy signal of output. Both AP and Lange assume that employer learning is independent of job assignment. In this paper, the variance of the initial error and the precision of the signal of output vary by initial occupational assignment. Throughout the analysis, I do not model initial occupational choice and assume that it is associated with a fixed learning parameter for the worker's entire career.[6] This is a limitation

pation A is twice the level in occupation B. Suppose that at the start of the career, employers in both occupations know nothing about their potential hires, so that the variance of pay is zero in occupation A and B. Assume that after five years employers in both occupations have learned half of what there is to know about the workers. In this example, the *growth* in the cross-sectional variance of pay will be larger in occupation A compared to occupation B, even if the speed of learning is the same across the two occupations. On the other extreme, occupations A and B can have the same level of productivity variance but employers in A learn about the worker's productivity in 2 years while completing the learning process takes 5 years in occupation B.

[4] Some of the relevant references are Topel and Ward (1992), Gibbons et. al (2005), Gibbons and Waldman (1999; 2006), Lemieux (2006), Lange (2007), among others.

[5] AP's learning model closely follows that of Farber and Gibbons (1996). One important difference between the two is that FB estimate a wage level regression while AP's dependent variable is the logarithm of wages.

[6] The idea here is that the initial occupation sets the worker on a specific career track. Since employer learning occurs early in the worker's experience profile, initial job assignment is likely to play the strongest role in revealing the worker's ability and thus affecting her subsequent occupational assignments.

since different occupational sorting patterns can generate differences in the level of cross-sectional productivity variance. In the empirical analysis I provide evidence that occupations with different learning parameters do not appear to have different underlying variance in the AFQT score.

I test two principal hypotheses. If differences in employer learning across initial occupational assignments are empirically significant, the growth in the AFQT coefficient and the growth in wage dispersion, the two measures of learning I use, will be similar within groups of occupations with similar employer learning patterns. In contrast, the two measures of learning will differ across occupations with different learning parameters.[7]

Empirically, I use data from the Current Population Survey (CPS) Outgoing Rotation Group (ORG) for the 1984-2000 period to calculate the two-digit occupation residual variance at each experience level in the first ten years of the worker's career.[8] I then calculate the *growth* in wage dispersion for workers at different stages in the experience profile.[9,10] I merge the estimated growth rates with the first occupation that workers report in the 1979 National Longitudinal Survey of Youth (NLSY79) and compare the growth in the AFQT coefficient

[7] The flow of information about the worker is public and observed by all labor market participants. The debate in the literature wether information flows symmetrically across employers has not been resolved yet. Recent papers such as Kahn (2007) and Schönberg (2007) use a learning model to test whether information between employers is symmetric. Although they use the same data set they reach opposite conclusions.

[8] I focus on the first 10 years in the worker's career since most of the employer learning occurs early in the experience profile. Lange estimates that employers' initial expectation errors are reduced by 50% within the first 3 years of the workers' career.

[9] Ideally, I would want to follow the growth in the variance of residual wages for a cohort of workers who started their career in the same occupation over time, rather than comparing the variance of two cohorts within an occupation. This is not feasible in the NLSY79 because of sample size limitations. In section 5, I perform robustness checks and provide evidence that occupational mobility is not the main mechanism behind the paper's results.

[10] The classification of occupations from the CPS based on the growth in the residual variance does not change based on the window I use to calculate the growth measure. That is, occupations which exhibit high growth in the residual variance between the second experience year and labor market entry are the same occupations that exhibit high growth in the residual variance between the fifth year of experience and labor market entry and the tenth year of experience and labor market entry.

across occupations with different growth rates in wage dispersion.[11]

My findings suggest that occupations with high growth in the variance of residual wages are also the occupations with high growth in the AFQT coefficient. I call occupations with high levels of growth in wage dispersion and in the AFQT coefficient "high learning" occupations. Interestingly, "low learning" occupations, where the growth in wage dispersion and in the AFQT coefficient is low, are not occupations where employers know little about the worker after 10 years of experience. The high AFQT coefficient at labor market entry in these occupations suggest that employers have already learned about the worker's AFQT score at the time of hire. I also show that my occupational classification into learning groups does not change significantly if the classification is based on the educational-specific growth in the residual variance, confirming the important implication of job assignment on the process of employer learning (Arcidiacono, 2010).

In order to test whether improved occupational matching rather than employer learning explains the results, I restrict my sample to workers who do not change their initial occupation after 5 and 10 years in the labor market and show that the results are not affected by this restriction. In a separate exercise, I limit my sample to initial occupations that are associated with low occupational mobility. Even in this limited sample, there appear to be significant differences in employer learning patterns across occupations.

The rest of the paper proceeds as follows. Section 2 presents a learning model which incorporates the idea of differential learning rates across jobs. Section 3 describes the data,

[11]Focusing on occupation-specific growth rates in the variance of residual wages compared to looking at levels of residual wages is important since it accounts for the fact that employers in some occupations might have richer information about worker's ability, that is not related to learning. It also accounts for the fact that levels of residual wages might be a mechanical consequence of how broadly or narrowly the occupation is defined and not of employer learning.

and section 4 includes the empirical analysis and results. Section 5 includes robustness checks and section 6 concludes and discusses future work.

# 2 Conceptual Framework

## 2.1 Sources of Employer Learning

Differences in the learning patterns across occupations can be driven by the *amount* to be learned in each occupation, because of different underlying differences in the productivity variance, or by differences in the *speed* of learning. A standard model of employer learning, such as the one formulated by AP, starts by specifying the log productivity of individual $i$ as a function of the information available to employers and researchers at every experience level and a function that captures the experience profile of productivity. In the standard model, log productivity does not depend on the worker's occupation. To allow for this possibility, assume that there are $J$ jobs (or occupations) in the economy.[12] For the purpose of this paper, I define an occupation as the entire expected career track associated with choosing a given initial occupation, so that a worker's entire career is associated with a fixed learning parameter. Worker $i$'s log productivity, $h$, at experience level $x$ in occupation $j$ is decomposed into four components

$$h_{ij,x} = as_{ij} + bq_{ij} + \Gamma z_{ij} + \eta_{ij} + \tilde{H}_j(x_i),\ j = 1, ..., J \tag{1}$$

[12] Throughout the paper I use the terms job type and occupation interchangeably. The empirical analysis will use occupations to test for differential employer learning.

The variables $(s, q, z, \eta)$ describe the different types of information available to employers and the econometrician in each occupation. In (1), $s$ represents variables observed by the employer and the econometrician (such as education); $q$ is a productivity component observed by the employer but is not in the data (such as information obtained during a recruitment interview); $z$ is a variable observed only by the econometrician (such as AFQT score); and $\eta$ is an individual productivity component that neither the employer nor the econometrician observe. $\tilde{H}_j(x_i)$ is a function that describes the experience profile of productivity in each occupation and is assumed to be independent from $s, q, z$,or $\eta$.[13]

Non-random occupational sorting (which I do not model in this paper) may generate differences in the cross-sectional productivity mean and variance across different occupations.[14] Variation in the underlying occupational productivity variance implies that the *amount* to be learned by employers might vary across occupations.

Because employers do not observe $z$ or $\eta$, they form expectations conditional on $s$ and $q$. In what follows, I suppress the index $i$. Assuming that these expectations are linear in both $s$ and $q$, define

$$\begin{aligned} z_j &= E(z_j|s_j, q_j) + \phi_j = \beta_1 q_j + \beta_2 s_j + \phi_j,\ j = 1, ..., J \qquad (2) \\ \eta_j &= E(\eta_j|s, q_j) + \varepsilon_j = \alpha_1 s_j + \alpha_2 q_j + \varepsilon_j,\ j = 1, ..., J \end{aligned}$$

By definition of an expectation, both $\phi$ and $\varepsilon$ are uncorrelated with $s$ and $q$ and have mean

[13] The focus in this paper is only on time-invariant productivity components, such as innate ability, and abstracts from time-variant components acquired, for example, from on-the-job training, although the two are likely to be correlated.

[14] For example, the productivity variance may be smaller in occupations attracting mainly high ability workers, such as engineers.

zero.[15] Calculating the expected value of the skill level in (1) conditional on the information available to employers upon entry to the labor market (at $x = 0$), and substituting (2) in (1) yields

$$h_j = (a + \Gamma\beta_2 + \alpha_1)s_j + (b + \Gamma\beta_1 + \alpha_2)q_j + (\Gamma\phi_j + \varepsilon_j) + \tilde{H}_j(x),\ j = 1, ..., J \tag{3}$$

The term $(\Gamma\phi_j + \varepsilon_j)$ in (3) represents the occupational-specific initial error in the employer's expectation about the worker's ability and is uncorrelated with $s$ and $q$.

After each period in the labor market, an occupation-specific noisy signal $y_{jt}$ of $h_j$ becomes available to all employers

$$y_{jt} = h_j + \nu_{jt}\ ,\ j = 1, ..., J \tag{4}$$

The noise $\nu_{jt}$ is job-specific and is uncorrelated with the other variables in the model or across initial occupations. As in other models of employer learning, $\nu_{jt}$ is assumed to be drawn from a normal distribution with variance $\sigma^2_{j\nu}$ and is assumed to be independently and identically distributed.[16] The worker's output history available to employers in occupation $j$ at each experience level $x$ is summarized by the vector $y_j^x = \{y_{j0}, y_{j1}, ..., y_{jx-1}\}$. Thus, the number of available productivity measures is equal to the experience of workers.

The *speed* of learning is the second potential source of difference in the learning process across occupations. In a seminal contribution, Lange, building on AP's model, formulated the conceptual framework to estimate the speed of learning. I adopt his formulation but

[15]I also assume that the error terms $\phi$ and $\epsilon$ are uncorrelated across occupations.

[16]In the empirical analysis I use the two-digit occupation codes to look at learning across jobs. This implicitly assume that the i.i.d shock, $\nu$, is the same across all the individual occupations under each code and does not take into account the fact that within-occupation job heterogeneity might be substantial.

allow the speed to vary across occupations. Lange shows that the posterior distribution at each experience level $x$ is normal where the mean $\mu_{jx}$ is

$$\mu_{jx} = (1 - \theta_{jx})E[h_j|s_j, q_j] + \theta_{jx}(\frac{1}{x}\sum_{t=0}^{x-1} y_{jt}) \tag{5}$$

and the precision is

$$\frac{1}{\sigma_{jx}^2} = \frac{1}{\sigma_{j0}^2} + \frac{x}{\sigma_{jv}^2} \tag{6}$$

Notice that $\sigma_{j0}^2$ is the variance of the initial expectation error $\Gamma\phi_j + \varepsilon_j$. This implies that employers in some occupations may predict better (or worse) workers' productivity compared to other occupations at the time of hire. The coefficient $\theta_{jx}$ at each experience level is given by $\theta_{jx} = \frac{xk_j}{1+(x-1)k_j}$. Lange refers to the parameter $k$ as the speed of employer learning where

$$k_j = \frac{\sigma_{j0}^2}{\sigma_{j0}^2 + \sigma_{jv}^2} \tag{7}$$

The speed of employer learning, $k$, is a function of the variance of the initial expectation error and the variance of the noise from the output signal. Employers in two occupations can have similar initial expectation error but differ in the variance of the output noise they receive each period. In this case, employers will learn faster about the worker's productivity in the occupation where $\sigma_{jv}^2$ is smaller at each experience level, and $\theta_{jx}$ will converge to 1 earlier in the experience profile. The speed of learning can be different, however, even if $\sigma_{jv}^2$ is similar across the two occupations and will be driven mainly by differences in $\sigma_{j0}^2$. In the extreme case where $\sigma_{j0}^2 = 0$, employers predict the worker's productivity at the time of hire and they do not use future outputs to update their initial assessment.

To complete the model I follow the standard assumption that wages are set by spot-market contracting, and at the end of each period workers are paid their expected output. That is, $W_j(s_j, q_j, y_j^x) = E[\exp(h_j)|s_j, q_j, y_j^x]$. The resulting log wage equation is[17]

$$w_j(s_j, q_j, y_j^x) = (1 - \theta_{jx})E[\tilde{h}_j|s_j, q_j] + \theta_{jx}(\frac{1}{x}\sum_{t=0}^{x-1} y_{jt}) + H_j(x) \tag{8}$$

Equation (8) describes the relationship between log wages, the information that employers have at the time of hire $(s_j, q_j)$, and the information that becomes available at each experience level.

## 2.2 Estimation Framework

Recall that only variables in $s$ and $z$ are observed in the data and can be used for the empirical analysis. Assuming that $s$ and $z$ are scalars, the log wage process from equation (8) can be summarized in the following regression

$$w_{jt} = \mu_{sjt}s + \mu_{zjt}z + H_j(t),\ t = 0, ..., T \tag{9}$$

The main insight of AP is that if employers learn about workers' productivity, the coefficient on a correlate of productivity which is not observed by the employer upon hiring, such as AFQT score, should increase over time.[18] We can estimate equation (9) by occupation and

[17] The conditional expected value of $\exp(h_j)$ is $(E[h_j|s_j, q_j, y_j^x] + \frac{1}{2}\sigma_{jx}^2) = \exp(E[\tilde{h}_j|s_j, q_j, y_j^x] + \tilde{H}_j(x) + \frac{1}{2}\sigma_{jx}^2)$. This is because the distribution of $h$ conditional on $(s_j, q_j, y_j^x)$ is normal. In the log wage equation I denote $H_j(x) = \tilde{H}_j(x) + \frac{1}{2}\sigma_{jx}^2$.

[18] The wage regressions in AP control for initial occupations fixed effects, but do not analyze the wage dynamics across occupations. The proof of the main proposition in AP is not affected by introducing different

compare the evolution of the hard-to-observe variable over time in order to test for differences in employer learning. However, because of sample size limitations in the NLSY this strategy is not feasible.

Instead, I use the implication that employer learning should increase the variance of residual wages over time. I calculate changes in the residual variance between different experience levels in each occupation and combine this information with the initial occupation reported by workers in the NLSY sample. I group occupations with similar growth patterns in the residual variance and test the hypothesis that occupations with the highest increase in the residual variance between two experience levels will also be occupations where the AFQT coefficient grows the most.[19]

The use of the *growth* in the variance of residual wages as a measure of employer learning is important since it accounts for *level* differences in the variance of residual wages across occupations which might reflect, among other things, differences in the initial error in the expectations that employers form (differences in $\sigma^2_{j0}$).[20] The analysis, however, cannot distinguish whether learning is different because of differences in the underlying productivity variance or because the variance of the noisy signal is different and I am unable to identify the parameter $k$ for each occupation.[21,22]

informational structure across occupations.

[19] Theoretically employer learning generates growth in wage dispersion. However, there are other reasons which can explain growth in wage dispersion within and across occupations. For example, the growth in the wage dispersion in traditionally unionized occupations might be small even if employer learning rates are high (Lemieux, 2006).

[20] Using the growth in the residual variance also accounts for the fact that levels of residual wages at different experience levels might be a mechanical consequence of how broadly or narrowly the occupation is defined and not of employer learning.

[21] Theoretically this might seem like a significant drawback. Empirically, however, I find evidence for learning only in some occupations while I find that there is no learning in others so distinguishing between the mechanisms behind these differences becomes less important.

[22] Riley (1979) tests for the screening role of education across occupations with endogenous sorting. Combining Riley's framework with AP's learning hypothesis might be a fruitful avenue to disentangle the mech-

One remaining related issue that needs to be discussed is *timing*. Two occupations might have the same growth in the residual variance between the time of hire and 10 years into the career, but in one occupation the residual variance stopped growing earlier than in the other. To address the timing issue I calculate the growth in the variance between different experience intervals and test the sensitivity of the results under different scenarios. The empirical analysis offers more details about this point.

# 3 Data

Most of the analysis in this paper uses data drawn from the NLSY79 covering the 1979-2000 period. The NLSY79 is a nationally representative sample of 12,686 men and women who were between the ages of 14 and 22 when they were first interviewed in 1979. The survey was conducted annually until 1994 and since then the participants are interviewed on a biennial basis. The NLSY79 has been used in many of the studies of employer learning because of two main features: detailed information on work experience allows the calculation of actual rather than potential labor market experience, and most importantly, the sample includes some variables that are correlated with ability but are not observed (or used) by the employer, such as AFQT scores and father's education.

The work history file in the NLSY79 provides information on the hours worked in each week of the year. In order to calculate actual labor market experience, I follow AP's methodology and accumulate the number of weeks in which the worker reports to have worked more than 30 hours and divide it by 50. I focus only on jobs after an individual has left school

anisms behind the differences in employer learning across occupations.

for the first time and drop all observations before that. The month and year when a person left school for the first time is directly reported by the respondents and is used as the entry date to the labor market. An individual can go back to school and remain in the sample.

To construct my sample, I include only white or black men which leaves me with 5,403 individuals. I drop 2,256 respondents who left school before 1979 because information on their first occupation and actual labor market experience is hard to determine. I drop 362 respondents who never report to have left school and 8 individuals who do not report the month they left school for the first time. 5 respondents who never completed more than 8 years of schooling are excluded from the sample, and 101 individuals who do not have a valid AFQT score are dropped. I finally drop 606 additional individuals who do not report their first occupation.

Aside from initial occupation, individuals in the NLSY79 report information on all the jobs they held between two interviews. I only use information from the job they hold at the time of the interview (CPS item). I exclude jobs without pay, jobs at home, and military jobs. The wage measure in the sample is the hourly wage rate of pay at the most recent job from the CPS section of the NLSY79. I use CPS deflators to calculate real wages in 1984 prices and I drop wages below \$1 or above \$100. Finally, because the linear specification is almost surely misspecified as we add more experience years (Lange, 2007), I restrict my sample to observations with 13 years of potential experience or less. This experience interval captures the approximately linear region in the relationship between log wages and the AFQT score over the experience profile (Arcidiacono, 2010). My final sample includes 18,700 observations on 2,065 distinct individuals.

Since NLSY79 participants took the AFQT at different ages, I standardize the AFQT scores to have mean zero and a standard deviation of 1. I do this by subtracting the mean score for a person of that age group and divide it by the standard deviation for that age group. Table appendix A1 provides summary statistics for the main variables I use in the empirical analysis.

In order to calculate growth in residual variance by occupation, I use information from the ORG CPS. The optimal way to calculate growth in the residual variance and attribute it to employer learning would be to track an individual at different points in the experience profile. Since a large panel that allows to follow individuals who had similar initial occupations over their career is not available I conduct a cohort-based analysis. Specifically, I calculate the residual variance at every experience level for the first 10 years of potential experience in each of the 1980 2-digit occupation codes. The ORG data set is adequate for the purposes of this paper because it contains information about the hourly wage rate, the same measure I use in the NLSY79 sample. Moreover, as Lemieux (2006) shows, the wage information in the ORG is less noisy compared to the March CPS because it measures directly the hourly wage of workers paid by the hour.

I pool data from 1984-2000 and keep in the sample only white or black men who have 0-10 years of potential experience.[23] I exclude from the sample self employed and full or part time students and individuals with below 8 years of schooling. I use the hourly wage rate for workers paid by the hour (about 60 percent of the sample) and calculate an hourly wage rate for the other workers by dividing their weekly earnings by their usual weekly hours. Real wages are calculated in 1984 prices using the CPS deflator. I drop wages that are below $1

[23] I analyzed other shorter periods of time such as 1984-1988 and 1992-2000. The results regarding the growth in the variance of wage residuals within occupations are not sensitive to the time period. In order to increase my sample size I choose to pool the entire period of 1984-2000.

and above \$100, and adjust top-coded earnings by a factor of 1.4 (Lemieux, 2006). Since actual experience is not reported in the ORG, I calculate a measure of potential experience (age-education-6). As it is well known, in 1992 the U. S. Bureau of Labor Statistics changed the educational attainment question in the CPS from one that was based on years spent in school to one that focuses on the highest degree received. I use the method proposed by Jaeger (1997; 2003) to create a consistent measure of years of schooling (to calculate potential experience) and to group individuals in consistent educational categories.[24]

# 4 Results

## 4.1 Ranking Occupations

I start the analysis by calculating the variance of residual wages by occupation for every experience level from the start of the worker's career up to 10 years of experience. I then calculate the growth in the variance of residual wages for three intervals: 0-5, 5-10, and 0-10 years of experience.[25] This choice of experience intervals should help me determine the importance of timing in the learning process.

Although this growth measure does not account for the possibility that occupational mobility might differ across initial occupations it has two main advantages. First, this

[24] I do not report in the paper descriptive statistics from the CPS sample. The distribution of educational groups across the sample period remains relatively constant with a steady increase throughout the period in the share of college graduates and those with some college. The distribution across experience groups is also similar across the sample years. This suggests that compositional effects over the sample period are small. I also computed the distribution of occupations over the sample period and find no evidence of substantial changes. Moreover, I analyzed the distribution of education by occupation-experience cells and found no evidence for significant changes in the education level within an occupation over the experience profile.

[25] I also calculated the growth in the residual variance between 0-2 years of experience. These alternative measures do not affect the ranking and classification of occupations.

growth measure does not confound the increase in wage dispersion due to learning with the increase in wage dispersion due to an increase in the price of unobserved skill because of skill-biased technological change (Katz and Murphy, 1992). Second, the growth measure will not be contaminated by potential composition effects due to large changes in the education and experience of the U.S. population (Lemieux, 2006; Card and DiNardo, 2002). In section 5, I address the concern regarding occupational mobility, and show that it is not what drives my results.

The residual wage is computed from a flexible regression of log hourly wage on an unrestricted set of dummies for age and years of schooling. I also include interaction terms between six schooling dummies and a quartic in age.[26] To capture any year-specific differences I include a full set of year dummies. I also include a set of occupation dummies to capture differences in the *level* of wage residuals between occupations. I weight the regression using weights provided by the CPS.

To calculate the variance in the residual wages, I use the 1980 2-digit occupation codes to create 45 occupation dummies and interact them with 10 experience cells. The coefficients of a regression of the squared residuals from the wage regression on the occupation-experience cells (450 coefficients) are the coefficients of interest. I rank occupations based on the total growth measure in the first 10 years (0-10). Ranking the occupations based on the growth measure of 0-5 years of experience produces similar occupational classification, with very few exceptions.

Table 1 provides a list of the occupations ranked by the growth in the variance of residual

[26] The six schooling dummies I use to interact with a quartic in age are 9-10, 11, 12, 13-15, 16, and 17+. I do not include an unrestricted set of age-education dummies because many of the cells are empty. I do not include other typical demographic variables like marital status and race in order to focus on direct measures of skill.

wages between 0-10 years of experience. The first six occupations, classified with the letter "H", have the highest total growth in the residual variance. Importantly, the residual variance in this group of occupations grows both in the first 5 years of experience and between 5-10 years. Occupations classified with the letter "M" have lower total growth in the residual variance between 0-10 years of experience.[27] For some of these occupations the growth in the residual variance is concentrated in the five first years of the experience profile (marked with a star). For others, the total growth in the residual variance in the first 10 years is split evenly between 0-5 and 5-10 years of experience (marked with 2 stars). Finally, for a small set of occupations, the growth in the variance between 0-10 years of experience is concentrated in the latter part of the experience profile. Thus, although the total growth in the variance in these occupations is similar in the first 10 years the timing of growth seems to be different. Occupations classified with the letter "L" have low growth in the residual wage in all experience intervals.

Notice that the ranking of occupations based on the growth in residual variance is different from ranking occupations based on mean education. For example, physicians have the highest growth in the variance while college professors experience a much lower growth, but both have high mean education. Nonetheless, some occupations with low or no growth in the residual variance appear to have high mean education, perhaps reflecting the fact that employers in these occupations have more information about their workers' abilities upon their hire. This is an important point since Arcidiacono et al. (2010) show that learning is mainly concentrated among high school graduates. I will show later in my robustness checks that

[27] The cutoff point was chosen such that the total growth in the residual variance is statistically different between the two occupations around the cutoff point. In this case, the growth of 0.15 in the sales workers, retail and personal services occupation is statistically different at the 5 percent level from the growth of 0.103 in the financial records occupation.

the ranking of occupations does not change if I calculate education specific growth in the residual variance in each occupation.

Table 1 also lists the means and standard deviation of the AFQT score in each initial occupations. As can be seen, although there are differences in the AFQT variance across occupations, there is no systematic difference in the standard deviation across the three learning groups. This suggests that differences in the underlying productivity variance due to endogenous sorting is not the main factor driving differences in learning across the three learning classifications. I provide more detailed evidence on this point later in the paper.

## 4.2 Evidence for Employer Learning

In this section, I reproduce AP's results using the updated sample. As discussed in section 2, the idea of testing for employer learning hinges on the availability of a variable which is (positively) correlated with ability, is unobserved by the employer at the time of hiring but is available to the econometrician. The empirical specification that stems from equation (9) is

$$\log W_{it} = \mu_{s0} s_i + \mu_{z0} z_i + \gamma T_i + \mu_{s10}(s_i \times \frac{T_i}{10}) + \mu_{z10}(z_i \times \frac{T_i}{10}) + \psi_{it} \tag{10}$$

Throughout the empirical analysis education ($s$) and AFQT ($z$) are interacted with experience divided by 10, so that the coefficients on the interaction terms represent the change in the wage slope between $T = 0$ and $T = 10$. Table appendix A2 reproduces AP's results using the original sample period of 1979-1992, and using the extended sample for the 1979-2000 period. I restrict the samples to include workers with 13 years of potential experience or less to capture the linear part of the learning process (Arcidiacono, 2010). All the reported standard errors are Huber-White standard errors. In the regression, experience is modeled with a cubic polynomial, and as in AP, I control for the 1980 two-digit occupation codes at the first job, urban residence, race, year fixed effects, and a linear trend. The base year for the time trend is 1992.[28]

Column 1 in Table A2 in the appendix reports the results of estimating equation (10) when the sample covers the 1979-1992 period and the interaction of AFQT with experience is excluded. The coefficient on education is 0.049 and statistically significant at the 1 percent level. The coefficient on the interaction term of education with experience is 0.022 and is also statistically significant. As in other studies that have used AFQT scores in wage regressions, it enters the equation with a positive (0.035) and statistically significant coefficient (Neal and Johnson, 1996). Column 2 uses the same restricted sample but includes an interaction term between AFQT and experience. As predicted by the model, the coefficient on AFQT when $T = 0$ becomes close to zero and statistically insignificant while the coefficient on the interaction term enters the regression with a large coefficient. At $T = 10$ the coefficient on AFQT is 0.086 and statistically significant at the 1 percent level. This indicates that employers do not observe the worker's AFQT score upon entry to the labor market but do learn about it over time. Extending the sample to cover the 1979-2000 period does not change the main results. As reported in column 3 and 4 in Table A2, the magnitude of the AFQT-experience coefficient at $T = 0$ is close to zero and statistically insignificant while the coefficient at $T = 10$ is 0.080 and statistically significant at the 1 percent level. In both columns 2 and 4, the coefficient on the education-experience interaction is close to zero and

[28]The time trend controls for economywide changes in the wage structure during the sample period (Katz and Murphy, 1992). I also experimented by adding a cubic time trend and interactions of the cubic time trend with education and race. These additional controls do not affect the results.

statistically insignificant.

### 4.3 Occupation-Specific Employer Learning

To test whether employer learning differs across occupations, I match the *first* occupation of each worker in the NLSY sample with the appropriate growth in the variance of residual wages of that occupation.[29] To start the analysis, I group occupations in three categories as described in Table 3. If the growth measure reflects (at least partially) employer learning, the growth in the AFQT coefficient should be increasing with the variance of residual wages. After 10 years of experience, we expect the highest growth in the AFQT coefficient in occupations with the highest residual variance growth rates.[30] In contrast, we expect little or no growth in the AFQT coefficient in occupations with low growth rates in the variance of residual wages. I will report results from running separate regressions for each of these groups, and then proceed by analyzing the sensitivity of the results to different alternative explanations. Notice that all models include occupation fixed effects, which among other things, also capture level differences in the residual variance upon the worker's hire.

Table 2 reports the results from estimating equation (10) for the three separate groups I describe above. The results in columns 1, 3, and 5 show significant heterogeneity in the

[29] About 80 percent of the workers in my sample change their initial 2-digit occupation after 10 years in the workforce. However, it may be limiting to consider only workers who never switch occupations. For example, an entry level sales position in a firm might be the starting point for many possible career paths in the company, and thus we would not want to restrict the sample to workers who stayed in sales. The real interest of the paper is in the entire menu of career paths that are associated with different initial occupational assignments. In the robustness checks I show that occupational mobility does not drive my results.

[30] Following the growth in the AFQT coefficient after 5 years in the labor marker (by interacting the AFQT coefficient with experience divided by 5 instead of 10) does not change the results qualitatively but simply cuts them in half. A more flexible approach would be to interact the AFQT score with dummy variables for each experience level. I tried to follow this approach but the small sample size within groups of occupations makes it hard to implement.

learning patterns across different groups of initial occupations. In column 1 I report the results on occupations with high growth in the residual variance, they constitute about 10 percent of the entire sample. The AFQT coefficient at $T = 0$ is negative -0.091 and statistically significant at the 1 percent level. This is different from AP's findings that the coefficient on the AFQT score should be zero at the time of hire. The growth in the coefficient, however, is consistent with the learning hypothesis. The coefficient on the AFQT score at $T = 10$ is 0.160 and statistically significant. Thus, occupations with the highest growth in the residual variance are also the occupations with the largest change in the AFQT coefficient over time. The coefficient on the education-experience interaction is 0.057 and significant at the 1 percent level. This is again different from the results in Table A2. One way to interpret the increase in the education coefficient over time is that high productivity workers in these occupations receive a considerable amount of on the job training over time. This, in turn, casts some doubt about interpreting the increase in the AFQT coefficient as evidence for employer learning since it might simply reflect on-the-job training. In the robustness checks I show that the AFQT coefficient for high school graduates in this group of occupations grows more than the AFQT coefficient for college graduates, supporting the learning perspective.

The results in column 3 refer to the group of occupations with intermediate growth in the residual variance and constitute about 79 percent of the sample. These occupations have learning patters which are consistent with the results in Table A2. The coefficient on the AFQT score at the time of hire is close to zero and is statistically insignificant while the coefficient on the interaction term of AFQT with experience is positive (0.082) and significant. As in the results for the entire sample, the returns to education do not change over time. The differences between the coefficients on the AFQT score at $T = 0$ and at $T = 10$ between column 1 and 3 are statistically significant at the 1 percent level. As mentioned earlier, although these occupations have similar total growth in the residual variance during the first 10 years, there are some differences in the timing of the growth of the residual variance.

To further explore how these differences affect the growth in the AFQT coefficient I disaggregated the intermediate learning occupation group into 3 subgroups. The first includes occupations where most of the growth in the residual variance is concentrated in the first 5 years (marked with a star), the second includes those where the total growth is split approximately equally across the two groups (marked with two stars), and the third includes occupations where the growth is concentrated in the latter part of the experience profile (not marked). The results among the three subgroups (not reported here) are similar to the results of the entire intermediate learning group and are not statistically different from each other.

I also tried instead of using a linear interaction of AFQT with experience to interact the AFQT with the 10 experience dummies and check whether the profiles are different among the three subgroup. The estimates from this exercise are not as precisely estimated but again do not reveal substantial differences in the AFQT-experience profile among the three subgroups. This should not be entirely surprising since, although I use the growth in the residual as an indicator for employer learning, it is likely to capture other factors which are not reflected in the growth of the AFQT coefficient but generate differences in the timing of the growth in the residual variance.

The results in column 5 refer to the group of occupations with low growth in the residual variance. As expected, the coefficient on the AFQT-experience interaction term is close to zero and statistically insignificant. This is an indication that employers in these occupations do not learn about worker's productivity. Of course, occupations can be low learning for two different reasons: First, it might be the case that the market has already learned the worker's AFQT at the start of the career; Second, it might mean that the market has not learned the worker's AFQT even after 10 years into the career. The large coefficient on AFQT at $T = 0$ in column 5 (0.051) supports the first scenario. It indicates that the worker's AFQT in low learning occupations is revealed immediately, or soon after their hire. This is plausible if employers in these occupations have access to richer information about workers' skills upon their hire (e.g. more information in $q$) which enables them to correctly estimate the worker's AFQT score.

To summarize, the results indicate considerable amount of heterogeneity in the learning process across different occupations. The AFQT coefficient at the time of hire increases monotonically as the growth in the residual variance decreases while the coefficient on the interaction term of AFQT with experience decreases monotonically as the growth in the residual variance decreases.[31]

As I mentioned earlier, I do not model in this paper initial occupational sorting and thus I am not able to distinguish whether learning differs across occupations because the underlying productivity variance is different or because the variance of the output signal that employers receive vary by occupation. I address this issue partially by calculating the mean and standard deviation of the AFQT score by occupation, and report them in the last

[31] I also tried to estimate the slope of the AFQT-experience profile for each learning group. Consistent with the results in Table 2, the slope of the AFQT-experience profile increases with the growth in the residual variance and is statistically different among the three learning groups.

two columns of Table 1. As can be seen, the mean and standard deviation of the AFQT score varies significantly across individual occupations, but not systematically across the different occupational learning groups. The mean AFQT score for all occupations in the high learning group is 0.313 with standard deviation of 0.927. The corresponding figures for the intermediate learning group is 0.009 and 0.981 while the mean and standard deviation for the low learning occupations are 0.490 and 1.003, respectively. Thus, although the level of mean AFQT score differs across the three groups, the standard deviation around that mean is similar. This observation suggests, although does not prove, that differences in the variance of the output signal play a bigger role in explaining differences in employer learning across occupations.

It is also interesting to look at the coefficient on the black dummy in the regressions across the different groups. Table 2 shows that the coefficient on the black dummy in high learning occupations is negative and large (-0.151), but is much smaller and insignificant (-0.022) for workers in low learning occupations (Column 5). An important insight of AP is that we can use learning models to test for statistical discrimination based on race. Although the focus of this paper is different, it is interesting to show how the coefficient on the race variable evolves over time in occupations with different patterns of employer learning.

The race variable can be treated as an $s$ variable or a $z$ variable. AP show that if there is a negative correlation between race and skill, and employers use it as a source of information about workers' skills, the coefficient on the black variable should be negative at $T = 0$. However, if employers initially statistically discriminate against black workers but learn over time, we would expect the racial gap to shrink, that is the coefficient on race to rise over time (decrease in absolute value). If employers, however, do not discriminate on the basis of race, the race variable can be treated as a $z$ variable. In this case, if race is negatively correlated with productivity and employers learn, the race gap will *widen* as workers accumulate experience.

The results in columns 2, 4, and 6 in Table 2 add to the regressions the interaction term of race with experience. The coefficient on the black dummy in column 2 is -0.171 and is statistically significant. The coefficient on the interaction term is positive (0.034) but is not precisely estimated. In contrast to AP the results suggest (although weakly) that employers in these occupations statistically discriminate on the basis of race. In occupations with intermediate learning patterns, the coefficient on the black dummy at the time of hire is -0.032 and increases (in absolute value) to -0.089 at $T = 10$. This is consistent with AP's findings that employers in these occupations do not discriminate on the basis of race. Lastly, the black coefficient in column 6, where there is no evidence for employer learning, is small, positive, and not statistically significant (0.014). Again, there is no evidence of statistical discrimination at the time of hire. The coefficient on the black-experience interaction in this group of occupations is negative (-0.064) but not statistically significant.

# 5 Robustness Check

## 5.1 Employer Learning Across Educational Groups

In a recent paper Arcidiacono et al. (2010) [hereinafter Arcidiacono] document different employer learning patterns across high school and college graduates. Their results suggest that while employers learn about the AFQT score of high school graduates, the AFQT coefficient for college graduates is positive and large at the time of hire and does not increase significantly over time. Are differences in employer learning across occupations simply a reflection of the underlying educational composition across occupations?

Columns 1 and 3 in panel A of Table 3 reproduce the results reported in Arcidiacono for high school and college graduates, respectively. The coefficient on the AFQT score for high school graduates is close to zero and statistically insignificant and, consistent with the learning hypothesis, the coefficient on the AFQT score after 10 years of experience is 0.062, significant at the 1 percent level. Adding initial occupational fixed effects in column 2 does not change the results significantly. This is not the case for college graduates. Consistent with Arcidiacono, the AFQT coefficient at the time of hire is 0.080 (column 3), significant at the 1 percent level while the coefficient on the AFQT score after 10 years of experience is small (0.037) and is not statistically significant. Adding occupational fixed effects, however, changes the results significantly. As can be seen in column 4 of Table 3 panel A, the AFQT coefficient at the time of hire becomes smaller although still positive and significant (0.028) and the coefficient on the AFQT score after 10 years becomes larger (0.043) and statistically significant at the 10 percent level.

To further explore whether my results are driven by the educational composition in each occupation, I recalculate the education-specific residual variance for every experience level in the first 10 years in the workers' career, by occupation.[32] Table appendix A3 reports the growth in the residual variance between the time of hire and 10 years after for high school and college graduates. Table A3 also reports the growth measure used in Table 2

[32] In the remainder of the analysis I focus only on high school and college graduates. In some occupations, there were not enough observations to calculate the residual variance at the different experience levels. I label these with NA = Not Available.

and the percentage of high school and college graduates in each occupation. Interestingly, the occupations I classify as high learning occupations in Table 2 are also the occupations with the highest growth in the residual variance for both high school and college graduates. Few of the occupations originally classified in the intermediate learning group would change classification depending on the growth of the residual in the two educational groups. For example, the growth in the residual variance for college graduates in financial records and processing occupations is 0.186 which would classify this occupation as high learning, while it is only 0.066 for high school graduates in the same occupation which would classify it as an intermediate learning occupation. There are 8 occupations out of 27 intermediate learning occupations where the classification would change by educational groups.

More significant differences appear among the group of occupations originally classified as low learning. For example, while the growth in the residual variance is 0.104 for high school technicians, except health engineering and science, it is 0.014 for college graduates in the same occupation. In fact, for five occupations out of 11 in this group the growth in the residual variance is higher for high school graduates than for college graduates.

Based on these findings, the main results of the paper in Table 2 should not change significantly for the high and intermediate learning occupations when split by educational group. In contrast, the results for the low learning occupations (no growth in the AFQT coefficient) should be reproduced only for college graduates.

Panel B of Table 3 reproduces the results of Table 2 using the original classification of occupations. The results for the high school graduates in column 1 are in line with the results of column 1 in Table 2. The AFQT coefficient at the time of hire is -0.096 while the coefficient on the AFQT-experience interaction is 0.318. Similarly, the AFQT coefficient for the college graduates in the same group of occupations is 0.008 and statistically insignificant and the coefficient on the AFQT-experience interaction is 0.065, although it is not precisely estimated. The results suggest that employer learning varies across occupations for both educational groups.

The results for the intermediate learning occupations are also consistent with the learning hypothesis for both groups of education. In fact, the growth in the AFQT coefficient for college graduates is bigger than the growth of the coefficient for the high school graduates (see column 3 and 4). As for the low learning occupations, the coefficient on the AFQT score is 0.066 for high school graduates at the time of hire and the coefficient on the interaction term is 0.098, both significant at the 10 percent level. This is different than the results reported in Table 2, column 6 where the coefficient on the interaction term is zero. These findings suggest that employers observe (at least to some extent) the high school graduates AFQT score but continue to learn about it over time. As for the college graduates in low learning occupations, the AFQT coefficient is large and significant at the time of hire (0.092) while the coefficient on the interaction term is negative but not statistically significant. This suggests that learning across occupations

In all occupational groups, except for the high school graduates in high learning occupations (column 1), the racial wage gap widens over time, regardless of educational attainment. Finally, reclassifying the occupations based on Table A3, where for some occupations the growth in the residual variance is different among high school and college graduates, and repeating the analysis produces stronger evidence of differences in the process of employer learning across occupations.[33]

[33] These results can be provided to the reader upon request.

A competing hypothesis to employer learning is OJT which can also generate increases in the residual variance over time (Ben-Porath, 1967). If AFQT and training are positively correlates, then the AFQT coefficient will also increase with experience. If the intensity of training differs by ability (regardless of the direction of the correlation) then the patterns across educational groups within groups of occupation should be different. The results, however, suggest that this is not the case. If, on the other hand, the intensity of training varies across the three occupational groups for both high school and college graduates then the current analysis cannot distinguish between OJT and employer learning. Collecting training measures at the occupational level can be a promising methodology to fully distinguish between the two hypotheses.

## 5.2 Occupational Mobility

Because of data limitations in the NLSY, the preceding analysis used the growth in wage dispersion using CPS data at different experience levels within an occupation. The theoretical model, however, implies that we need to track the growth in wage dispersion for workers who *start* their career in occupation $j$ over their experience profile. Combining the cross-sectional measures of variance with longitudinal data from the NLSY can be flawed if occupational mobility in the NLSY is substantial. Moreover, if initial occupational assignments are associated with different mobility patterns, the growth in the AFQT coefficient over time might reflect improved occupational matching rather than employer learning.

In order to address this issue I generate a dummy variable that takes a value of 1 if the worker reported to be in the same occupation five years into his career as when he was first hired. On the aggregate level, the share of workers who do not switch occupations after five years in the labor market is 21, 18, and 28 percent in the high, intermediate, and low learning occupations, respectively. The high rates of mobility of young workers are not surprising and in line with the findings of Topel and Ward (1992). It is encouraging, however, that mobility patterns are not substantially different across the three occupational categories.

To check the sensitivity of the results, Table 4 reproduces the results across the three occupational categories when limited only to workers who did not switch initial occupations after 5 years in the labor market. The results are encouraging as they portray similar patterns reported for the whole sample. Repeating the analysis while restricting the sample for workers who did not switch initial occupations even after 10 years in the labor market also produces similar results but with less precision due to smaller sample size. I also experimented by limiting the samples to initial occupations where 25 percent or more workers did not switch occupations after 5 years in the labor market. The results remain in line with the main findings but are not reported in the paper.

As a final attempt to address this issue, I calculated two measures of growth-in-dispersion using NLSY data. The first compares the wage dispersion of different experience cohorts within an occupation (similar to what I did using the CPS data), and the other compares the wage dispersion of workers who start their career at the same occupation, with their *own* wage dispersion after 5 and 10 years of experience (regardless of the current occupation in which they are in).[34] I, then, restrict my sample to occupations for which the two measure are *not* different from each other. There are 19 such occupations (out of 45), covering 49 percent

[34] Due to sample size limitation, I was not able to calculate the residual variance at every experience level. Instead, I aggregat workers with 0-5 years of experience and workers with 6-10 years of experience.

of my original sample size. Examples of such occupations are engineers, administrative occupations, construction trades, health technicians, and farm workers. As in the preceding analysis, this exercise reproduces the main results of the paper indicating that the findings I report are not driven by occupations for which the wage dispersion of workers over time differs from the dispersion of cohorts within an occupation.[35]

## 5.3 Alternative Occupational Ranking

As I show in Table 1 the ranking of occupations, with very few exceptions, remains the same whether I use the growth in the residual variance between the time of hire and 10 years into the worker's career or the growth in the residual variance in the first 5 years of the career. In particular few occupations that were ranked in the intermediate learning group would be classified as high learning occupations (such as farm operators and managers and teacher, college and university). Reclassifying these few occupations and repeating the analysis does not affect the main conclusion that learning patterns vary across groups of initial occupations.

I also experimented by calculating the growth in the residual variance in the first 2 years of the worker's career and ranking the occupations based on this measure. With very few exceptions, the ranking of occupations and their classification into the three groups of occupations does not change.[36] This gives me confidence that the paper's results are not sensitive to the growth measure I use to classify occupations into learning groups.

[35] These results were reported in an earlier version of the paper and can be provided to the reader upon request.

[36] Obviously, the value of the growth measure and the cut off points between the three groups of occupations change but the classification of occupations into one of the three groups remains intact.

# 6 Conclusion

In this paper I provide evidence that employer learning about the worker's AFQT score varies across initial occupational assignments. Employer learning has two principal implications which I utilize in the analysis. First, the wage dispersion of a cohort of workers increases as they gain experience. Second, the coefficient on a correlate of ability that employers initially do not observe, such as the AFQT score, grows with experience. If employers' learning varies significantly across occupations, both of these indicators of learning should vary across groups defined by a worker's initial occupational assignment.

This paper tests this implication of the employer learning model. I estimate an occupation-specific growth rate in the variance of residual wages for workers with 0-5, 5-10 and 0-10 years of experience using ORG CPS data for the 1984-2000 period, and match these values with data from the NLSY79. The results suggest that occupations with high growth in the variance of residual wages over the first 10 years are also the occupations with high growth in the AFQT coefficient. I also find that the AFQT coefficient in occupations with low growth in the variance of residual wages is large and does not grow over time. This indicates that employers in these occupations have already learned the worker's AFQT at the start of the career.

There are two different reasons why employer learning varies across occupations. Endogenous occupational sorting can generate differences in the underlying productivity variance which will affect the amount to be learned by employers. Two occupations, however, may have the same underlying productivity variance but differ in the speed of learning. In this paper, I am unable to disentangle the contribution of these two mechanisms, and this point awaits future research.

Testing the employer learning hypothesis by occupations is also a promising framework to disentangle the learning hypothesis from two competing hypotheses, namely improved occupational matching and on-the-job training. Finally, differences in learning across occupations highlights the importance of private information that workers might have about their own ability, and how it might affect their educational choices and occupational sorting.

# References


[1] Altonji, Joseph G. 2005. Employer learning, statistical discrimination and occupational attainment. *American Economics Review*, *Papers and Proceedings* 96:112-17.

[2] Altonji, Joseph G., and Charles Pierret. 2001. Employer learning and statistical discrimination. *Quarterly Journal of Economics* 116:313-50.

[3] Antonovics, Kate, and Limor Golan. 2007. Experimentation and job choice. Working Paper.

[4] Arcidiacono Peter, Patrick Bayer, and Aurel Hizmo. 2010. Beyond signaling and human capital: education and the revelation of ability. Forthcoming in *American Economic Journal: Applied Microeconomics*.

[5] Ben-Porath, Yoram. 1967. The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75:352-65.

[6] Card, David, and John E. DiNardo. 2002. Skill-biased technological change and rising wage inequality: some problems and puzzles. *Journal of Labor Economics* 20:733-83.

[7] Farber, Henry S., and Robert Gibbons. 1996. Learning and wage dynamics. *Quarterly Journal of Economics* 111:1007-47.

[8] Gibbons Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent 2005. Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics* 23:681-723.

[9] Gibbons Robert, and Michael Waldman. 1999. A theory of wage and promotion dynamics inside firms. *Quarterly Journal of Economics* 114:1321-58.

[10] —. 2006. Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics* 24:59-107.

[11] Jaeger, David.A. 1997. Reconciling the new Census Bureau education questions: Recommendations for researchers. *Journal of Business and Economic Statistics* 15:300-09.

[12] —. 2003. Estimating the returns to education using the newest Current Population Survey education questions. *Economic Letters* 78:385-94.

[13] Jovanovic, Boyan. 1979a. Job matching and the theory of turnover. *Journal of Political Economy* 87:972-90.

[14] kahn, Lisa B. 2007. Asymmetric information between employers. Harvard University, mimeo.

[15] katz, Lawrence F., and Kevin M. Murphy. 1992. Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics* 107:35-78.

[16] Lange, Fabian. 2007. The speed of employer learning. *Journal of Labor Economics* 25:1-35.

[17] Lemieux, Thomas. 2006. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review* 96:461-98.

[18] Miller, Robert A. 1984. Job matching and occupational choice. *Journal of Political Economy* 92:1086-1120.

[19] Neal, Derek A., and William R. Johnson. 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104:869-95.

[20] Neal, Derek A., and Sherwin Rosen. 1998. Theories of the distribution of labor earnings. NBER Working Paper no. 6378, National Bureau of Economic Research, Cambridge, MA.

[21] Riley, John G. 1979. Testing the educational screening hypothesis. *Journal of Political Economy* 87:S227-S251.

[22] Schönberg, Uta. 2007. Testing for asymmetric employer learning. *Journal of Labor Economics* 25:651-92.

[23] Spence, Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87:355-79.

[24] Topel, Robert, and Michael Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107:441-79.

Table 1: Ranking of 1980 2-digit occupation codes by the growth in the variance of wage residuals ORG CPS sample-1984-2000

| Occupation | GRV+ 0-5 | GRV 5-10 | GRV 0-10 | Classification+ | Mean Education | Mean AFQT NLSY | SD# AFQT NLSY |
|---|---|---|---|---|---|---|---|
| Health diagnosing occupations | 0.156 | 0.135 | 0.291 | H | 17.85 | 1.134 | 0.449 |
| Forestry and fishing occupations | 0.118 | 0.166 | 0.284 | H | 12.17 | -0.477 | 0.657 |
| Personal service occupations | 0.028 | 0.117 | 0.145 | H | 13.23 | -0.167 | 1.099 |
| Sales representatives, finance, and business service | 0.077 | 0.053 | 0.130 | H | 15.20 | 0.171 | 1.131 |
| Secretaries, stenographers, and typists | 0.025 | 0.105 | 0.130 | H | 13.97 | 0.909 | 0.686 |
| Sales workers, retail and personal services | 0.114 | 0.037 | 0.150 | H | 13.15 | 0.396 | 0.773 |
| Financial records, processing occupations | 0.119 | -0.015 | 0.103 | M* | 14.03 | 0.413 | 1.164 |
| Food service occupations | 0.075 | 0.025 | 0.100 | M* | 12.47 | 0.025 | 1.004 |
| Supervisors and proprietors, sales occupations | 0.075 | 0.022 | 0.097 | M* | 13.91 | 0.574 | 0.966 |
| Computer equipment operators | 0.062 | 0.025 | 0.100 | M* | 13.91 | NA | NA |
| Sales representatives, commodities, except retail | 0.039 | 0.039 | 0.078 | M** | 14.86 | 0.887 | 0.000 |
| Farm workers and related occupations | 0.046 | 0.031 | 0.077 | M** | 12.27 | -0.207 | 0.930 |
| Farm operators and managers | 0.170 | -0.097 | 0.073 | M* | 13.54 | 1.415 | 0.466 |
| Freight, stock and material handlers | 0.055 | 0.018 | 0.072 | M* | 12.25 | -0.422 | 1.004 |
| Supervisors - administrative support | 0.069 | 0.004 | 0.072 | M* | 14.39 | 0.536 | 1.145 |
| Other administrative support occupations | 0.046 | 0.022 | 0.068 | M* | 13.54 | 0.051 | 0.914 |
| Other handlers, equipment cleaners and laborers | 0.032 | 0.036 | 0.068 | M** | 12.07 | -0.320 | 0.906 |
| Motor vehicle operators | 0.044 | 0.023 | 0.067 | M* | 12.47 | 0.057 | 0.961 |
| Teachers, except college and university | 0.056 | 0.011 | 0.067 | M* | 16.22 | NA | NA |
| Other executive, administrators, and managers | 0.034 | 0.023 | 0.057 | M** | 15.00 | 0.650 | 0.753 |
| Protective service occupations | 0.028 | 0.028 | 0.056 | M** | 13.53 | -0.215 | 0.970 |
| Construction trades | 0.028 | 0.026 | 0.055 | M** | 12.32 | -0.170 | 0.902 |

+GRV=Growth in residual variance between different experience intervals. H=high learning; M=Intermediate learning. L=low learning. #SD=standard deviation

Table 1 continue: Ranking of 1980 2-digit occupation codes by the growth in the variance of wage residuals ORG CPS sample-1984-2000

| Occupation | GRV+ 0-5 | GRV 5-10 | GRV 0-10 | Classification+ | Mean education | Mean AFQT NLSY | SD# AFQT NLSY |
|---|---|---|---|---|---|---|---|
| Mechanics and repairers | 0.042 | 0.013 | 0.054 | M* | 12.65 | -0.188 | 0.732 |
| Cleaning and building service occupations | 0.044 | 0.010 | 0.054 | M* | 12.25 | -0.001 | 0.993 |
| Machine operators and tenders, except precision | 0.029 | 0.023 | 0.052 | M** | 12.19 | -0.332 | 0.953 |
| Health technologists and technicians | 0.010 | 0.037 | 0.047 | M | 14.35 | 0.704 | 0.842 |
| Construction laborer | 0.027 | 0.013 | 0.041 | M | 12.07 | -0.287 | 0.840 |
| Teachers, college and university | 0.127 | -0.088 | 0.040 | M* | 17.17 | 0.712 | 0.631 |
| Health service occupations | 0.003 | 0.033 | 0.037 | M | 13.04 | -0.117 | 0.844 |
| Engineering and science technicians | 0.007 | 0.024 | 0.032 | M | 13.94 | 0.893 | 0.663 |
| Engineers | 0.011 | 0.018 | 0.029 | M | 16.13 | 0.427 | 0.890 |
| Fabricators, assemblers, inspectors, and samplers | 0.007 | 0.018 | 0.026 | M | 12.34 | 0.176 | 0.958 |
| Management related occupations | 0.010 | 0.014 | 0.023 | M | 15.79 | 0.959 | 0.688 |
| Mail and message distributing | 0.051 | -0.037 | 0.014 | L | 13.09 | -0.337 | 1.013 |
| Other precision production occupations | 0.003 | 0.009 | 0.012 | L | 12.71 | 0.015 | 0.918 |
| Technicians, except health engineering and science | 0.003 | 0.006 | 0.009 | L | 15.34 | 1.066 | 0.774 |
| Mathematical and computer scientists | -0.008 | 0.010 | 0.002 | L | 16.01 | 1.474 | 0.000 |
| Private household service occupations | -0.019 | 0.016 | -0.002 | L | 12.34 | 1.004 | 0.768 |
| Lawyers and judges | -0.131 | 0.015 | -0.116 | L | 17.81 | 1.419 | 0.586 |
| Professional specialty occupations | -0.058 | 0.023 | -0.035 | L | 15.64 | 0.291 | 1.004 |
| Health assessment and treating occupations | -0.021 | -0.038 | -0.059 | L | 15.83 | 0.081 | 1.218 |
| Natural scientists | 0.012 | -0.076 | -0.064 | L | 16.72 | 0.559 | 0.691 |
| Administrators and officials, public administration | -0.099 | 0.010 | -0.088 | L | 16.02 | -0.252 | 0.993 |
| Transportation occupations and material moving | -0.028 | 0.022 | -0.006 | L | 12.21 | 0.076 | 0.805 |

+GRV=Growth in residual variance between different experience intervals. H=high learning; M=intermediate learning L=low learning. #SD=standard deviation

Table 2: The effect of AFQT on log wages by growth rates in the variance of wage residuals

| | High growth in wage dispersion | | Intermediate growth in wage dispersion | | Low growth in wage dispersion | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| AFQT | -0.0912* | -0.0948* | 0.0011 | 0.0126 | 0.0511* | 0.0568* |
| | (0.028) | (0.030) | (0.007) | (0.008) | (0.019) | (0.020) |
| AFQT*exp/10 | 0.1607* | 0.1684* | 0.0824* | 0.0597* | 0.0100 | -0.0004 |
| | (0.040) | (0.047) | (0.011) | (0.013) | (0.029) | (0.031) |
| Education | 0.0394* | 0.0397* | 0.0558* | 0.0551* | 0.0691* | 0.0682* |
| | (0.011) | (0.011) | (0.003) | (0.003) | (0.009) | (0.009) |
| Education*exp/10 | 0.0565* | 0.0558* | 0.0045 | 0.0057 | -0.0221 | -0.0201 |
| | (0.016) | (0.017) | (0.005) | (0.005) | (0.020) | (0.012) |
| Black | -0.1516* | -0.1706* | -0.0783* | -0.0320** | -0.0215 | 0.0143 |
| | (0.032) | (0.053) | (0.008) | (0.014) | (0.021) | (0.038) |
| Black*exp/10 | | 0.0369 | | -0.0887* | | -0.0640 |
| | | (0.082) | | (0.023) | | (0.056) |
| Share of Black | | 0.23 | | 0.28 | | 0.18 |
| $R^2$ | 0.37 | 0.37 | 0.37 | 0.37 | 0.51 | 0.51 |
| N | 1,824 | 1,824 | 14,754 | 14,754 | 2,122 | 2,122 |

*,** indicates significance at the 1and 5% level, respectively. The dependent variable is log real hourly wage. The experience measure is actual experience and is modeled as a cubic polynomial. All models control for the two-digit first occupations, urban residence, year effects, and a time trend with base year 1992. All standard errors are Huber-White standard errors

Table 3: The effect of AFQT on wages by education-specific occupational groups.

| Panel A | High School | | | | College | |
|---|---|---|---|---|---|---|
| | (1) | (2) | | | (3) | (4) |
| AFQT | 0.0175 | 0.0116 | | | 0.0802* | 0.0284*** |
| | (0.010) | (0.010) | | | (0.013) | (0.018) |
| AFQT*exp/10 | 0.0623* | 0.0609* | | | 0.0375 | 0.0438*** |
| | (0.018) | (0.018) | | | (0.025) | (0.025) |
| Black | -0.0408** | -0.0311 | | | 0.1240* | 0.0139 |
| | (0.019) | (0.019) | | | (0.041) | (0.034) |
| Black*exp/10 | -0.0970* | -0.0971* | | | -0.1787* | -0.1111** |
| | (0.031) | (0.031) | | | (0.059) | (0.050) |
| Initial occupation | NO | Yes | | | NO | Yes |
| $R^2$ | 0.22 | 0.25 | | | 0.22 | 0.33 |
| N | 6,933 | 6,649 | | | 4,510 | 4,219 |

| Panel B | High learning | | Intermediate learning | | Low learning | |
|---|---|---|---|---|---|---|
| | HS | Collge | HS | Collge | HS | Collge |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| AFQT | -0.0957** | 0.0076 | 0.0123 | 0.0089 | 0.0658*** | 0.0921* |
| | (0.050) | (0.051) | (0.011) | (0.022) | (0.036) | (0.033) |
| AFQT*exp/10 | 0.3179* | 0.0650 | 0.0429** | 0.0776** | 0.0975*** | -0.0740 |
| | (0.087) | (0.083) | (0.019) | (0.032) | (0.058) | (0.048) |
| Black | -0.0568 | 0.1024 | -0.0322 | -0.0316 | -0.0529 | 0.1280** |
| | (0.106) | (0.084) | (0.021) | (0.043) | (0.074) | (0.065) |
| Black*exp/10 | 0.1416 | -0.1350 | -0.1051* | -0.0899 | -0.1374 | -0.2457* |
| | (0.182) | (0.117) | (0.033) | (0.063) | (0.111) | (0.091) |
| $R^2$ | 0.24 | 0.31 | 0.24 | 0.29 | 0.32 | 0.31 |
| N | 327 | 636 | 7,726 | 2,916 | 546 | 649 |

*,**,*** indicates significance at the 1%, 5%, and 10% level. See list of controls in Table 4.

Table 4: The effect of AFQT on log wages by growth rates in the variance of wage residuals for workers with low occupational mobility

| | High growth in wage dispersion | Intermediate growth in wage dispersion | Low growth in wage dispersion |
|---|---|---|---|
| | (1) | (3) | (5) |
| AFQT | -0.2580* | 0.0020 | 0.0463*** |
| | (0.061) | (0.019) | (0.021) |
| AFQT*exp/10 | 0.3353* | 0.1093* | 0.0126 |
| | (0.101) | (0.026) | (0.034) |
| $R^2$ | 0.45 | 0.45 | 0.45 |
| N | 390 | 2,596 | 588 |

*,**,*** indicates significance at the 1%, 5%, and 10% level. See list of controls in Table 4.

# APPENDIX

Table A1: Descriptive Statistics, NLSY79 sample, 1979-2000

| Variable | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| Real hourly wage | 8.67 | 5.78 | 3 | 99.38 |
| Log of real hourly wage | 2.01 | 0.51 | 1.09 | 4.60 |
| Education | 13.81 | 2.45 | 8 | 20 |
| Black dummy | 0.26 | 0.44 | 0 | 1 |
| Actual experience | 5.38 | 3.50 | 0 | 16.26 |
| Standardized AFQT score | 0.095 | 0.99 | -2.99 | 2.41 |

The sample includes 18700 observations from 2,065 individuals.

Table A2: The effect of AFQT on log wages

| Model | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Education | 0.0488* | 0.0561* | 0.0475* | 0.0560* |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Black | -0.0731* | -0.0719* | -0.0821* | -0.0804* |
| | (0.008) | (0.008) | (0.008) | (0.008) |
| Stan AFQT | 0.0354* | -0.0022 | 0.0397* | -0.0016 |
| | (0.004) | (0.007) | (0.004) | (0.006) |
| Education*exp/10 | 0.0220* | -0.0062 | 0.0219* | 0.0062 |
| | (0.005) | (0.005) | (0.004) | (0.005) |
| Stan AFQT*exp/10 | | 0.0856* | | 0.0798* |
| | | (0.012) | | (0.010) |
| Sample period | 1979-1992 | 1979-1992 | 1979-2000 | 1979-2000 |
| $R^2$ | 0.38 | 0.38 | 0.41 | 0.41 |
| N | 15,714 | 15,714 | 18,700 | 18,700 |

* indicates significance at the 1% level. The dependent variable is log real hourly wage. The experience measure is actual experience and is modeled as a cubic polynomial. All models control for the two-digit first occupations, urban residence, year effects, and a time trend with base year 1992. All standard errors are Huber-White standard errors.

Table A3: Ranking of 1980 2-digit occupation codes by the growth in the variance of wage residuals ORG CPS sample-1984-2000

| Occupation | GRV* High school | GRV College | GRV all | % High school | % College |
|---|---|---|---|---|---|
| Health diagnosing occupations | 0.170 | 0.288 | 0.291 | 0.54 | 98.19 |
| Forestry and fishing occupations | 0.301 | 0.221 | 0.284 | 57.64 | 7.22 |
| Personal service occupations | 0.113 | 0.456 | 0.145 | 39.31 | 18.19 |
| Sales representatives, finance, and business Service | 0.107 | 0.149 | 0.130 | 12.03 | 66.51 |
| Secretaries, stenographers, and typists | 0.113 | 0.148 | 0.130 | 28.70 | 30.13 |
| Sales workers, retail and personal services | 0.126 | 0.219 | 0.150 | 41.50 | 17.73 |
| Financial records, processing occupations | 0.066 | 0.186 | 0.103 | 25.54 | 32.46 |
| Food service occupations | 0.091 | 0.145 | 0.100 | 48.93 | 8.73 |
| Supervisors and proprietors, sales occupations | 0.090 | 0.120 | 0.097 | 32.91 | 33.68 |
| Computer equipment operators | 0.100 | 0.092 | 0.100 | 29.16 | 28.08 |
| Sales representatives, commodities, except retail | 0.120 | 0.104 | 0.078 | 18.39 | 59.42 |
| Farm workers and related occupations | 0.087 | 0.140 | 0.077 | 52.62 | 6.77 |
| Farm operators and managers | 0.102 | 0.071 | 0.073 | 41.81 | 28.97 |
| Freight, stock and material handlers | 0.082 | 0.074 | 0.072 | 59.81 | 4.29 |
| Supervisors - administrative support | -0.013 | 0.108 | 0.072 | 24.34 | 42.06 |
| Other administrative support occupations | 0.053 | 0.088 | 0.068 | 37.01 | 25.75 |
| Other handlers, equipment cleaners and laborers | 0.083 | 0.015 | 0.068 | 60.68 | 3.52 |
| Motor vehicle operators | 0.058 | 0.108 | 0.067 | 57.29 | 6.97 |
| Teachers, except college and university | 0.113 | 0.130 | 0.067 | 7.26 | 78.32 |
| Other executive, administrators, and managers | 0.108 | 0.042 | 0.057 | 17.53 | 58.15 |
| Protective service occupations | 0.071 | 0.136 | 0.056 | 60.64 | 5.30 |
| Construction trades | 0.062 | 0.008 | 0.055 | 57.53 | 5.71 |

*GRV=Growth in residual variance between different experience intervals

+H=high learning; M=medium learning; L=low learning

Table A3 continue: Ranking of 1980 2-digit occupation codes by the growth in the variance of wage residuals ORG CPS sample, 1984-2000.

| Occupation | GRV* High school | GRV College | GRV all | % High school | % College |
|---|---|---|---|---|---|
| Mechanics and repairers | 0.057 | 0.097 | 0.054 | 54.02 | 6.96 |
| Cleaning and building service occupations | 0.081 | 0.025 | 0.054 | 58.96 | 5.66 |
| Machine operators and tenders, except precision | 0.056 | 0.014 | 0.052 | 63.13 | 3.96 |
| Health technologists and technicians | 0.105 | 0.106 | 0.047 | 18.12 | 35.28 |
| Construction laborer | 0.039 | 0.111 | 0.041 | 57.44 | 4.35 |
| Teachers, college and university | 0.124 | 0.075 | 0.040 | 0.76 | 93.63 |
| Heath service occupations | 0.049 | 0.025 | 0.037 | 42.26 | 13.93 |
| Engineering and science technicians | 0.018 | 0.054 | 0.032 | 24.84 | 25.76 |
| Engineers | 0.038 | 0.023 | 0.029 | 4.39 | 86.45 |
| Fabricators, assemblers, inspectors, and samplers | 0.021 | 0.026 | 0.026 | 34.33 | 21.71 |
| Management related occupations | 0.020 | 0.042 | 0.023 | 6.38 | 80.14 |
| Mail and message distributing | 0.038 | -0.036 | 0.014 | 44.10 | 14.03 |
| Other precision production occupations | 0.037 | 0.010 | 0.012 | 54.81 | 10.57 |
| Technicians, except health engineering and science | 0.104 | 0.014 | 0.009 | 10.11 | 64.56 |
| Mathematical and computer scientists | 0.065 | -0.011 | 0.002 | 3.51 | 81.74 |
| Private household service occupations | -0.218 | 0.018 | -0.002 | 3.14 | 86.30 |
| Lawyers and judges | NA | -0.098 | -0.006 | 0.28 | 98.65 |
| Professional specialty occupations | -0.028 | -0.016 | -0.035 | 9.21 | 70.10 |
| Health assessment and treating occupations | 0.126 | -0.068 | -0.059 | 5.20 | 69.21 |
| Natural scientists | NA | -0.069 | -0.064 | 1.53 | 92.74 |
| Administrators and officials, public administration | NA | -0.011 | -0.088 | 7.69 | 51.75 |
| Transportation occupations and material moving | -0.029 | NA | -0.116 | 64.12 | 2.56 |

*GRV=Growth in residual variance between different experience intervals

+H=high learning; M=medium learning; L=low learning. NA=not available.