Testing for Asymmetric Employer Learning and Statistical Discrimination

Suqin Ge, Andrea Moro, and Beibei Zhu*

September 7, 2018

Abstract

We test the implications of a statistical discrimination model with asymmetric learning. Firms receive signals of productivity over time and may use race to infer worker's productivity. Incumbent employers have more information about workers productivity than outside employers. Using data from the NLSY79, we find evidence of asymmetric learning. In addition, employers statistically discriminate against non-college educated black workers at time of hiring. We also find that employers directly observe most of the productivity of college graduates at hiring, and learn very little over time about these workers.

Keywords: statistical discrimination, employer learning, asymmetric learning
JEL code: J71, D82, J31

1 Introduction

This paper develops and empirically tests the implications of asymmetric employer learning in a model with statistical discrimination. A crucial assumption in statistical discrimination models is that the productivity and qualifications of labor force participants is difficult to observe directly, therefore imperfectly informed employers use demographic characteristics such as rage or gender as proxies for unobserved worker characteristics.[1]

*Ge: Department of Economics, Virginia Tech. Moro : Department of Economics, Vanderbilt University. Zhu: Amazon. This study benefited from funding from the U.S. Department of Labor. The views expressed are those of the authors and should not be attributed to the Federal Government, the Department of Labor, or Amazon. Emails: ges@vt.edu, andrea@andreamoro.net, zhubeibei0319@gmail.com.

[1]There are two main branches of statistical discrimination theories, screening discrimination and rational stereotyping. The former, originated from Phelps (1972), attributes discriminatory outcomes to unexplained exogenous differences between groups, combined with employers' imperfect

In an influential paper, Altonji and Pierret (2001) (AP hereafter) adopted the landmark model of employer learning by Farber and Gibbons (1996) to derive testable implications of the statistical discrimination hypothesis. In their framework, employers learn about workers' productivity over time through newly acquired information such as job performance. A direct implication of the employer learning theory is that firms become less inclined to statistically discriminate based on observed group characteristics as they accumulate more information about individual workers' productivity. Hence, over time employers rely less and less on race as a proxy for productivity and wages become more correlated with measures of productivity available to the investigator.

Many subsequent studies provide further empirical evidence showing that employers learn over time (e.g., Lange (2007), Arcidiacono et al. (2010), and Mansour (2012)), but all of these studies are carried out under the assumption that employer learning is symmetric, that is, both incumbent and outside firms have the same information about workers' productivity.[2] On the other hand, a number of theoretical articles have considered the hypothesis that current employers are at an informational advantage about workers' productivity than outside employers, a phenomenon labeled as "asymmetric employer learning."[3] The nature of employer learning may have a large influence on the effects of statistical discrimination. If employer learning is purely asymmetric, that is, if the new employer learns nothing from a worker's previous jobs, members from the discriminated group would suffer on average a wage loss each time a new job is started, because learning has to restart at each job turnover. On the contrary, under symmetric learning the learning process continues regardless of job turnovers, thus members of the discriminated group suffer a wage penalty only at their first job. Our study develops a testable model that nests both symmetric and asymmetric learning hypotheses and provides new tests for statistical discrimination based on race.

information about workers' productivity. This literature (see also Aigner and Cain (1977)), is largely agnostic on where the initial group differences originate. They may result from either differences in employer perceptions or other factors, such as differences in the quality of education or human capital acquisition. The other branch of this theory, originated from Arrow (1973) and modeled most comprehensively by Coate and Loury (1993), assumes that employer's negative beliefs about the quality of minority workers are self-fulfilling and thus average group differences are endogenously derived in equilibrium. The other type of discrimination theory is taste-based Becker (1971), where employers have prejudice against minority workers. Fang and Moro (2011) provide a detailed survey on the theoretical literature on statistical discrimination, and Lang and Lehmann (2012) offer an extensive survey on theory and empirics of racial discrimination.

[2] Most of the studies focus on males using U.S. data. A notable exception is Lesner (2018), who finds evidence of statistical discrimination against women using a Danish sample.

[3] Examples of theoretical models of asymmetric employer learning include Waldman (1984), Greenwald (1986), Bernhardt (1995), and Golan (2005).

We propose a learning-based statistical discrimination model based on the seminal model by Phelps (1972). Employers have incomplete information and use group membership to infer workers' productivity; moreover, they use productivity signals over time to better assess workers' productivity. Statistical discrimination based on group membership arises because employers perceive average productivity differences between groups. We extend this model to allow the employer learning process to be asymmetric, that is, we assume that outside firms have less information about workers' productivity than current employers.

We show that asymmetric learning implies different predictions about how wages evolve with job experience versus job tenure. Symmetric learning implies a continuous learning process over a worker's general experience regardless of job turnovers. If outside firms have no information on a worker's productivity, learning will be interrupted once the worker moves from one job to another. Therefore, the learning process takes place over job tenure rather than general experience. As a consequence, the correlation of wages with measures of skills observed by the econometrician should increase more with tenure than with experience, a testable implication (the opposite occurs when learning is symmetric). In addition, when employers statistically discriminate against minorities, initially wages of minority workers (conditional on skills) are lower, but over time, as employers learn about productivity, the effects of race decreases conditional on measures of skills available to the econometricians. Finally, we show that asymmetric learning has implications regarding job mobility when employers statistically discriminate. The initial, negative effect of statistical discrimination is reset every time a minority worker changes jobs. Hence we should expect fewer job transitions among minority workers, conditional on their skills. This is not the case if learning is symmetric, because new employers inherit the information acquired by previous employers, therefore there are no negative consequences of changing jobs arising from statistical discrimination

We test these implications using the National Longitudinal Survey of Youth 1979 (NLSY79), the same data used in AP, but including more recent waves. We follow the literature in using the standardized value of the Armed Forces Qualification Test (AFQT), a battery of aptitude tests, as the measure of skills observed by the econometrician. Using the sample of non-college educated workers, we find evidence that employers learn asymmetrically about workers' skills and that they statistically discriminate against black workers. Wages become more correlated with skills as time passes, and this correlation increases more when tenure is used as a measure of time, as opposed to experience, consistently with asymmetric learning. Moreover, black workers without a college education suffer a wage penalty initially, but wages become more correlated with skills over time. In addition, and confirming the asymmetric learning hypothesis, we find lower mobility of non-college educated black workers conditional on AFQT.

Results are different for college educated workers. For this class of workers we find neither evidence that learning is asymmetric, nor evidence of statistical discrimination against african-american workers. We conjecture that key aspects of worker productivity are directly observed by employers upon initial labor market entry and thus little learning takes place subsequently, consistently with the main findings reported in Arcidiacono et al. (2010).

In addition to the papers we already cited that belong to the employer learning and statistical discrimination literature, our paper mostly relates to the literature testing for asymmetric learning. This empirical literature offers no conclusive evidence on the nature of employer learning. Schönberg (2007) studies a two-period model of asymmetric learning and derives implications for job transitions and wage dynamics, but ignores statistical discrimination. Using a sample of white males only, she finds that employer learning is mostly symmetric. Pinkston (2009) also tests the implications of asymmetric learning. In his model however, outside employers learn the information of current employers through competitive bidding; therefore, *employment spells*, as opposed to tenure or experience, is the relevant time variable in his analysis. His empirical results suggest that asymmetric employer learning plays a role that is at least as important as symmetric learning during an employment spell. Kahn (2013) also investigates asymmetric learning, using an original approach that looks at the implications on the variance of the wage changes for workers that change jobs and workers that stay with their current jobs. She finds support for asymmetric information between incumbent and external employers, but does not focus on racial differences. Several studies (Gibbons and Katz (1991), Bauer and Haisken-DeNew (2001), and Pinkston (2009)) find empirical evidence in favor of asymmetric employer learning, that is, current firms have access to more information about workers' productivity than outside firms do. Our contribution, relative to this literature, is to focus on the implications of asymmetric learning on racial differences generated by statistical discrimination, and to test them by educational level.

The structure of the paper is as follows. Section 2 describes the learning-based racial statistical discrimination model, and derives empirically testable predictions. Section 3 provides an overview of the data and compares the results in AP with the results from our sample using the same specification. The main results are reported in sections 4 and 5. Section 6 concludes suggesting directions for further research.

2 Theoretical Framework and its Empirical Implications

2.1 The model

We consider a signal extraction model where firms have more precise information about the workers they employ than about workers employed by other firms. To discuss the empirical implications of such asymmetry, we extend the standard statistical discrimination model in Phelps (1972) to include two types of employers and a time dimension to allow for employers' learning.

Firms compete for workers and maximize output given wages. Workers care only about wages. When they match with a firm, they draw a match-specific productivity q from a Normal distribution with mean $\mu(X)$ and variance $\sigma^2(X)$, where X is a set of variables observed by the employer that correlate with productivity. In the standard statistical discrimination model, X includes group identity such as race or gender. Ability and wages are expressed in logarithms to guarantee that they have positive values in levels.

Employers initially observe a signal of productivity s_0 from an unemployed worker, and then, after hiring, in each period t they observe from the employee an additional signal $s_t = q + \epsilon_t$, where ϵ_t is distributed normally with mean 0 and variance σ_ϵ^2. The signal's variance can be interpreted as a measure of the signal's information quality (higher variance corresponds to poorer quality). We assume that competition for workers drives wages to the value of workers' expected productivity.[4]

It is helpful to start the analysis by exploring the effect of the incumbent employer's learning. In each period, the employer computes the worker's expected productivity given the signals observed. New workers are offered expected productivity $E(q|s_0)$. Standard properties of the bivariate normal distribution[5] imply that

$$E(q|s_0) = (1-\alpha)\mu(X) + \alpha s_0,$$
$$\alpha = \frac{\sigma^2}{\sigma_\epsilon^2 + \sigma^2}.$$

In this expression, expected ability is a weighted average of the population average skill and the signal, with weights equal to the relative variance of the two variables. When the signal is perfectly informative ($\sigma_\epsilon = 0$), the population mean is ignored; when the signal is pure noise ($\sigma_\epsilon = \infty$), expected ability is equal to the population's average, conditional on observables X. With a partially informative signal,

[4]As we discuss below, this assumption relies on the inability of the incumbent employer to extract informational rents from the worker in the case of asymmetric learning

[5]See Eaton (1983).

the conditional expected ability is increasing in both q and s_1. The conditional distribution, which we denote with $\phi(q|s_1)$ is also normal, with mean equal to $E(q|s_1)$ and variance $\alpha\sigma_\epsilon^2$.

As the worker increases her tenure with a firm, the incumbent employer exploits information from multiple signals, which together provide more precise information about true productivity. Again exploiting normality, one can derive that after T periods (and $T+1$ signals) the expected productivity is:[6]

$$E(q|s_0, ..., s_T) = (1-\alpha_T)\mu(X) + \alpha_T \left(\frac{\sum_{t=0}^{T} s_t}{T+1} \right), \tag{1}$$

where:

$$\alpha_T = \frac{(T+1)\alpha}{1+T\alpha}. \tag{2}$$

Note that the weight α_T placed on the signals average is increasing in T and converges to 1. As tenure increases, the worker's expected productivity gets closer to her true productivity.

Consider an econometrician observing wages and a one-time signal of skill r not observed by the employer, such as the AFQT, such that $r = q + \epsilon_r$, $\epsilon_r \sim N(0, \sigma_r^2)$. Then,

$$cov(r, E(q|s_0...s_T)) = cov\left(q+\epsilon_r, (1-\alpha_T)\mu(X) + \alpha_T \left(\frac{\sum_t (q+\epsilon_t)}{T}\right)\right) = \alpha_T Var(q),$$

that is, as the number of signals increase, the expected productivity covaries more and more with the signal observed by the econometrician.

Empirical Implication 1. Under the assumption of the model, in a wage regression the interaction of workers' tenure with AFQT displays a positive coefficient.

The result does not rely necessarily on assuming perfect competition in the labor market. For example, if employers and workers bargain over a share of the (expected) surplus, a sufficient condition for the the implication to hold is that the bargaining power of the worker does not change with tenure too much.[7]

Consider now workers hired after being separated from their employers where they worked for $T-1$ periods. In period T,assume that workers draw a new match-specific productivity q^o from the same distribution, and a signal v for the new employer.

[6]See the appendix for details.

[7]Note that the model relies on final output being not contractible, which is the case for example when workers' contribution to total output cannot be observed with certainty.

We assume that new employers form priors that depend on available information from workers' curriculum and other signals, which we summarize with variable v. The expected productivity given this prior information is also normal, conditional on the worker willing to move, therefore:

$$E(q^o|v) = (1-\alpha^o)\mu(X) + \alpha^o v.$$

The workers' previous experience and job offer may be considered by the new employer when computing the conditional expectation of the worker's productivity, but this information enters only the available signal $v$with variance σ_ν^2. Crucially, we assume that the information available is worse than the information available to the current employer if she had stayed an extra year, that is:

$$\alpha^o < \alpha_T, \tag{3}$$

where T is the number of years the worker has been with the current employer.

If the new employer does not infer any information from prior job history, then her signal ν carries the same information as any other signal available to new employers, or $\sigma_\nu = \sigma_\epsilon$,, implying $\alpha^0 = \alpha$, which is less than α_T from equation (2), a situation we label as *purely asymmetric learning.* In the more general case, we believe assumption in (3) to be realistic because a worker's curriculum, job interviews, aptitude tests cannot substitute from day-to-day interactions over the workers' tenure.-[8]

After being hired, employers' expectations evolve in the same way they do for new hires, that is, employers extract a signal every period from the same distribution as the previoius employers' signals , and revise their posterior expectations using Bayes' rule. The formula after T periods from the job change is different because the first draw of v has a different variance (σ_ν^2) than subsequent draws (σ_ϵ^2). After the new hire the conditional distribution has variance $\frac{\sigma^2\sigma_\nu^2}{\sigma^2+\sigma_\nu^2}$, that is,

$$E(q|\nu, s_1, ...s_T) = E(1-\alpha_T^o)\left((1-\alpha^o)\mu(X) + \alpha^o\nu\right) + \alpha_T^o\left(\frac{\sum_{t=1}^T s_t}{T}\right),$$

[8] In a model where outside employer bid with current employers for workers' wages in an auction, Pinkston (2009) proves that the incumbent employer's information about workers is completely revealed to the outside employer after the bidding process. For our purposes, we assume that frictions exist in the environment that prevent the information to be completely revealed to competing employers, or that unemployment spells prevent the bidding process to completely reveal the information.

where

$$\alpha_1^o = \frac{\alpha^o \sigma_\nu^2}{\alpha^o \sigma_\nu^2 + \sigma_\epsilon^2} \text{ and } \alpha_T^o = \frac{T\alpha_1^o}{1+(T-1)\alpha_1^o}, \text{for } T > 1$$

(see the appendix for a derivation). If the signal available to the new employer is at least as good as the first signal available to the first employer (s_0),$\sigma_v^2 < \sigma_\epsilon^2$, which together with $\alpha^o < \alpha$, as assumed above, implies $\alpha_1^o < \alpha_1$ and in general $\alpha_T^o < \alpha_{T.}$

Compare for the sake of example, two workers, Mary and John, with the same experience $T+Q$. Mary has been always with the same employer, whereas John has worked for two employers, changing job after T periods, and staying for Q periods with the new employer. Then the expected productivities for the two workers can be compared by "breaking" the sequence of $T+Q$ signals observed by Mary's employer into two subsequences:

$$\begin{aligned}
\text{Mary:} \qquad & E(q|s_0, ...s_T) = E(1-\alpha_Q)\left((1-\alpha_T)\mu(X) + \alpha_T\left(\frac{\sum_{t=0}^{T} s_t}{T+1}\right)\right) \\
& \qquad + \alpha_Q\left(\frac{\sum_{t=T+1}^{T+Q} s_t}{Q}\right) \\
\text{John:} \qquad & E(q|\nu, s_{T+1}, ...s_{T+Q}) = E(1-\alpha_Q^o)\left((1-\alpha^o)\mu(X) + \alpha^o\nu\right) \\
& \qquad + \alpha_Q^o\left(\frac{\sum_{t=T+1}^{T+Q} s_t}{Q}\right)
\end{aligned}$$

It is now easier to see that the weight placed on the conditional mean $\mu(X)$ is larger for the worker who changed employer, John, because $\alpha^o < \alpha^T$ and $\alpha_Q^o < \alpha_Q$

This implies that wages of workers with discontinuous work histories covary with the econometricians' signals of productivity less than workers with continuous work histories, leading to the following:

Empirical Implication 2. Consider two regressions that include the interactions of either tenure with AFQT, or experience with AFQT, and assume that learning is asymmetric. Then, the coefficient of the interaction of tenure with AFQT is positive and larger than the coefficient of the interaction of experience with AFQT.

Some workers with high experience have low tenure, therefore the correlation of experience with AFQT is lower. Note that the opposite implication would be true if learning was symmetric. In that case, some workers with low tenure have high experience. Their employers had the opportunity to learn more, therefore the coefficient on tenure should be attenuated relative to the coefficient on experience.

We now extend the model to study its implications on statistical discrimination. There are two groups of workers with easily recognizable traits: a minority (M)

and a dominant (D) group. Assume that $\mu(M) < \mu(D)$ and that employers use race for labor market decisions.[9] Signals of productivity that are observed by both the econometrician and the employer are accounted for by the term $\mu(X)$, therefore variables in X will be over time less correlated with wages:

Empirical Implication 3. Under the assumptions of the model, in a wage regression including a race M dummy, if group M is statistically discriminated against, the coefficient on such dummy is negative, but its interaction with tenure is positive so that the negative effect declines over time.

Finally, we want to look at how incomplete information affects workers' job mobility. First, we note that under asymmetric learning (assumption 3) changing jobs corresponds, on average, to resetting one's wage closer to the population mean. Hence workers with higher skills should be less likely to change jobs.

Empirical Implication 4. Under asymmetric learning, workers with higher skills are less likely to change jobs than workers with lower skills.

If learning is symmetric instead, the new employers inheriting information from previous employers, there are no implications on job mobility across workers of different skills.

Next, we consider the implications on job mobility under statistical discrimination. Consider the choice of two workers with the same productivity q from groups D and M who have the option choosing to move or not. We maintain the assumption that $\mu(M) < \mu(D)$. Initially, the worker from group M is paid less than the worker from D, but over time both of their wages converge to their true productivity q.

Suppose the workers randomly receive offers from competing firms each drawing a match-specific productivity q' from a Normal distribution centered around q and with variance $\sigma^2(X)$. The new firm observes a noisy signal of q', $q' + \epsilon$, where as before $\epsilon \sim N(0, \sigma^2_\epsilon)$. Two forces play a role in making the probability of a job change higher for the worker from D. First, note that an attractive offer can only come from high q' draws. The worker may accept offers when $q' < q$ only if the initial signal ϵ is high enough to compensate the fact that eventually, the wage will converge to q'. Hence, on average, workers are less likely to accept an offer when q' is low, because they need a higher initial signal draw. Because the productivity distribution for D-workers first-order stochastically dominates the productivity distribution of M workers, for any q, the worker from group D is more likely to draw higher and

[9] We focus here on the empirical implications of such behavior ignoring its legal aspects: using race even for informational purposes is in general illegal, but employers may be able to do so by using other proxies for race. Ultimately, whether or not employers statistically discriminate is an empirical question.

acceptable productivity/signal combinations than the worker from group M. Second, for any acceptable productivity draw, initially the offered wage for M workers is lower than the initial offer for D workers because the new employer is initially putting a larger weight on $\mu(X)$. For a better intuition, consider the case where the signal for the outside firm is very noisy (α^o arbitrarily close to zero), whereas the worker, having worked for a long time for the current firm, receives a wage arbitrarily close to $q > \mu(X)$. The initial offer is very close to $\mu(D)$ or $\mu(M)$ depending on group identity, therefore the signal draw that is necessary to attract an M worker is higher than the signal draw that is necessary to attract a D worker.[10] Hence, workers from group D will be more likely to change jobs. We can state the following:

Empirical Implication 5. If group M is statistically discriminated against, conditional on AFQT group M workers are less likely to change jobs than workers from group D.

2.2 Empirical Specification

In this subsection we propose an empirical specification motivated by the theoretical framework presented above. We showed that if employer learning is symmetric, the learning process occurs over general work experience regardless of job turnovers. In contrast, purely asymmetric learning implies that only current employers learn about workers' productivity over time, so that learning only takes place over job tenure. To distinguish the two learning hypotheses, we use actual work experience X and job tenure T as two separate time measures. The corresponding wage equations we estimate are as follows:[11]

$$\begin{aligned}\ln w_i = {} & \beta_0^X + \beta_S^X S_i + \beta_{S,X}^X (S_i \times X_i) + \beta_{AFQT}^X AFQT_i + \beta_{AFQT,X}^X (AFQT_i \times X_i) \\ & + \beta_{Black}^X Black_i + \beta_{Black,X}^X (Black_i \times X_i) + \beta_\Omega^X \Omega_i + H(X_i) + \epsilon_i^X, \end{aligned} \tag{4}$$

$$\begin{aligned}\ln w_i = {} & \beta_0^T + \beta_S^T S_i + \beta_{S,T}^T (S_i \times T_i) + \beta_{AFQT}^T AFQT_i + \beta_{AFQT,T}^T (AFQT_i \times T_i) \\ & + \beta_{Black}^T Black_i + \beta_{Black,T}^T (Black_i \times T_i) + \beta_\Omega^T \Omega_i + H(X_i) + \epsilon_i^T. \end{aligned} \tag{5}$$

where w_i denotes the hourly wage of individual i, S_i measures years of schooling, $AFQT_i$ denotes individual AFQT score, $Black_i$ is a dummy variable on race, and Ω_i

[10] Note that since q and q' are independent draws from the same distribution, there is nothing to be learned by the employer from the fact that the worker accepts, hence the new employer's expected productivity does not need be to be conditioned on the fact that the worker accepts.

[11] In our model, learning is nonlinear in time, which implies that the effects of AFQT score and race should also vary nonlinearly with time. For simplicity, however, we follow the literature and assume the relationships between log wage, AFQT score, and race to be linear in time.

is a vector of demographic variables and other controls. In all of our specifications, we control for urban residence, dummies for region of residence, and year fixed effects. The variables X_i and T_i measure time, and $H(X_i)$ is a polynomial in experience. Time is measured in months in our sample, and we divide the interaction of any variable with time measure by 120 so the coefficients on interaction terms measure the change in wage during a ten-year period. In the empirical analysis below we follow the literature and assume the effects of AFQT and *Black* on log wages to vary linearly with time to simplify the interpretation of these coefficients.

Empirical implication 1 suggests that the coefficient $\beta^T_{AFQT,T}$ should be positive. By implication 2, it should be significantly larger than $\beta^T_{AFQT,X}$. Under statistical discrimination, implication 3 suggests $\beta^T_{Black} < 0$ and $\beta^T_{Black} > 0$, and $\beta^X_{Black} < 0$ and $\beta^X_{Black} > 0$, when learning is asymmetric.

To test empirical implications 4 and 5, we specify the following probit model that examines the effect of race on workers' probability of job change:

$$Pr(J_{i,t} = 1) = \Phi(\beta_0 + \beta_1 AFQT + \beta_2 Black_i + \beta_\Omega \Omega_{i,t}), \tag{6}$$

where $J_{i,t}$ is a dummy variable for job change, with $J_{i,t} = 1$ if individual i changes job in month t and $J_{i,t} = 0$ if he stays on the same job; $Black_i$ is a dummy variable on race; and $\Omega_{i,t}$ is a vector of individual demographic and other control variables. Our theoretical analysis suggests that under asymmetric learning we should observe $\beta_1 < 0$ and, when employers statistically discriminate, $\beta_2 < 0$.

3 Data

The empirical analysis is based on the 2008 release of NLSY79, a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994 and on a biennial basis thereafter. These data contain detailed information on family background, academic performance and labor market outcomes of a cohort of young workers, and its weekly work history data provide information to construct accurate measures of actual work experience and job tenure.

The empirical analysis is restricted to black and white male workers who have completed at least eight years of education, thus we use the same restriction AP. We only analyze labor market observations after a person makes school-to-work transition. An individual is considered to have entered the labor market when he leaves school for the first time. Following the criteria used in Arcidiacono et al. (2010), military jobs, self-employed jobs, jobs at home, and jobs without pay are excluded from the construction of experience and from the analysis as we want to focus our analysis on civilian employees.

We construct individual monthly employment status using NLSY79 work history data, which contains each respondent's week-by-week labor force status since January 1978. An individual is considered as employed in a given month and accumulates one month of work experience or tenure if he works at least 10 hours per week for at least three weeks, or during the last two weeks in the month. Otherwise, an individual is classified as nonemployed. The work history information is employer-based, thus a "job" should be understood as an uninterrupted employment spell with an employer. We link all jobs across survey years and build a complete employment history for each respondent in the sample. Multiple jobs held contemporaneously are treated as a new job, with an associated wage equal to the average wage weighted by hours on each job, and working hours equal to the sum of working hours on the different jobs. Tenure on a job is completed when an individual makes a job-to-job transition or whens she is back in non-employment. Job tenure is the number of months between the start of a job and either the date the job ends or the interview date. Actual work experience is the sum of tenure for all jobs.[12] Potential work experience is defined as months since the respondent first left school.

The wage measure that we use is the hourly rate of pay on each job, provided in the work history file. Nominal wages are deflated to real hourly wages in 1990 dollars by using the monthly CPI released by the BLS. Real wages less than \$1 or more than \$100 per hour are excluded from the analysis. We use the AFQT as our proxy correlate of productivity. To eliminate age effects, we standardize the AFQT score to have a mean zero and standard deviation one for each three-month age cohort. We use data from both the main cross-sectional sample of the NLSY79 and the supplementary sample, which oversamples blacks and disadvantaged whites.[13] The total remaining sample consists of 2,595 whites and 1,136 blacks with 320,124 monthly observations.

We also consider in the analysis two education samples: white or black men who have completed at least 16 years of education (college graduates sample) or less than 16 years of education (non-college graduates sample).[14]

[12] In AP, actual experience is defined as the weeks worked divided by 50. Our measure is very close to theirs and more compatible with our tenure measure.

[13] All statistics in this study are unweighted. Using sampling weights does not change the qualitative results.

[14] We have experimented with restricting the sample to those who have exactly a high school or a college degree following Arcidiacono, Bayer and Hizmo (2010). The results are similar. As both high school dropouts and workers with some college education but without a college degree behave similarly to high school graduates, we bundle them into a sample of workers with no college degree.

Table 1: Summary Statistics by Race

	Whites			Blacks		
	All	<College	≥College	All	<College	≥College
AFQT	0.501	0.201	1.345	-0.571	-0.726	0.482
	(0.957)	(0.885)	(0.568)	(0.796)	(0.649)	(0.897)
Education (yrs)	13.35	12.14	16.74	12.69	12.12	16.60
	(2.39)	(1.31)	(1.19)	(2.00)	(1.35)	(1.06)
Hourly wage	12.91	11.10	17.99	10.15	9.23	16.41
	(8.14)	(5.89)	(10.97)	(6.14)	(4.98)	(8.96)
Experience:						
Potential	131.41	135.07	122.12	145.37	147.71	129.46
	(84.51)	(86.20)	(78.66)	(85.76)	(86.50)	(78.73)
Actual	110.54	111.97	106.52	111.15	111.28	110.25
	(76.05)	(77.50)	(71.68)	(73.17)	(73.66)	(69.76)
Job tenure	46.94	45.98	49.63	40.90	40.17	45.92
	(48.42)	(48.29)	(48.70)	(43.30)	(42.76)	(46.50)
Individuals	2,592	1,906	686	1,133	987	146
Observations	224,304	165,480	58,824	93,684	81,660	12,024

Notes: Standard deviations are in parentheses. Education is measured in years, real hourly wages in 1990 dollars, and experience in months. Potential experience is months since left school.

Table 1 presents the summary statistics for the main variables in our sample by race and education level. The average AFQT score of black workers is about one standard deviation lower than that of white workers, possibly as a result of pre-market discrimination or racial bias in testing. This test score gap persists even if we control for education. Black workers generally earn lower wages and accumulate less job tenure than white workers. Potential employers have strong incentives to statistically discriminate on the basis of race if AFQT is a good measure of skill. In the next section, we carry out the empirical analysis to examine this issue in detail.

In Table 2, we compare the results from AP's specification using different samples. We report for convenienence in Column (1) the results from AP's Table 1, Panel 1, Column 4. The specification in column (2) uses data from the same time period (interview years 1979–1992), but with several differences in sample construction that we adopt for our analysis. First, we use monthly data instead of annual data. Secondly, we measure potential experience as time since first left school instead of age minus years of schooling minus 6.

Despite the differences, the main qualitative results from AP are confirmed. The coefficient on education is positive and significant initially and falls over time. The

Table 2: Sample Comparisons

	(1)	(2)	(3)	(4)
Time	1979–1992	1979–1992	1979–2008	1979-2008
Education	0.079***	0.088***	0.071***	0.080***
	(0.015)	(0.007)	(0.006)	(0.006)
Education × experience/120	-0.019	-0.035***	-0.002	-0.018*
	(0.013)	(0.009)	(0.004)	(0.007)
Standardized AFQT	0.022	0.035*	0.057***	0.036**
	(0.042)	(0.014)	(0.012)	(0.013)
AFQT × experience/120	0.052	0.069***	0.037***	0.071***
	(0.034)	(0.018)	(0.009)	(0.014)
Black	-0.057	-0.030	-0.037	-0.039
	(0.072)	(0.026)	(0.022)	(0.025)
Black × experience/120	-0.083	-0.084**	-0.053***	-0.053*
	(0.058)	(0.031)	(0.015)	(0.026)
R^2	0.287	0.273	0.346	0.322
No. of Observations	21,058	177,288	317,988	212,640

Notes: Column (1) reproduced from AP's Table 1, Panel 1, Column 4. In columns (2)-(4), the experience measure is months since left school for the first time. All specifications control for year effects, urban residence, region of residence, experience, and experience squared. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

coefficients on AFQT and AFQT–experience interaction imply that the impact of AFQT score on log wages rises as workers accumulate more experience. That is, employers learn about workers' productivity over time, so the weight they put on the hard-to-observe correlate of productivity, AFQT, increases. The coefficient on *Black* is small and insignificant at the time of initial hire, but it becomes significantly negative over time. As the racial wage gap is initially not statistically different from zero, AP conclude that there is no statistical discrimination on the basis of race.

Column (3) reports analogous results using our full sample, the 1979-2008 waves of NLSY79. We obtain qualitatively similar results, but AFQT and *Black* now have flatter profiles with experience and the returns to AFQT are greater initially. The difference in the time paths of AFQT and *Black* is likely driven by a non-linear employer learning process. In order to make our sample more comparable to the AP sample, in Column (4) we restrict our sample to experience level less than 168 months, the maximum months of potential experience in the AP sample. This restriction restores the lower initial AFQT effect and its steeper profile over time.

4 Results

An important finding in the employer learning literature is that the employer learning process may vary across different educational groups. Arcidiacono et al. (2010) find that a college degree helps workers directly reveal key aspects of their productivity, and thus employer learning is more important for high school graduates. They argue that if all education levels are pooled in wage regressions, the estimates can be biased and the results may be misinterpreted. Based on their results, we split our sample into college graduates and non-college graduates, where a person who has completed at least 16 years of education is considered a college graduate and otherwise a non-college graduate.[15] We use these two samples from NLSY 1979–2008 to test the main predictions of our learning-based statistical discrimination model.

4.1 Wages

Our empirical analysis has two main focal points. First, to distinguish between symmetric and asymmetric employer learning, we examine the initial coefficients on AFQT as well as their interaction terms with time when experience and tenure are used as two separate time measures in log wage equations (4) and (5). Secondly,

[15] Arcidiacono et al. (2010) restrict their sample to white or black men who have exactly a high school or a college degree with 12 or 16 years of education. If we restrict our college sample to those with 16 years of education and our high school sample to those with 12 years of education, the empirical results are very similar to those we find below.

we investigate how the racial wage gap varies over time to examine whether or not employers statistically discriminate against black workers. If employers hold racial prejudice, our learning-based statistical discrimination model predicts a large initial racial wage gap because employers base payments on race and a narrowing racial gap over time as the employers accumulation more information on true productivity.

In Panel 1 of Table 3 we report estimates of the wage regressions using the non-college graduate sample. If employer learning is symmetric, learning takes place over general work experience. The specification in column (1) estimates equation (4) with actual work experience in months as the experience measure. We use actual work experience because it is a more accurate measure of workers' labor market experience than potential experience and the construction of actual experience and tenure are more consistent with each other. Actual experience is determined by workers' employment decision, which is correlated with individual productivity. This unobserved heterogeneity across individuals may produce inconsistent estimates of the effect of experience on wages as well as the speed of employer learning over experience. In addition, actual experience may be used by employers as a measure of quality (it is an indicator of the intensity of worker effort). Because of these potential endogeneity concerns, we include in column (2) the results from an Instrumental Variables (IV) specification where actual experience is instrumented with potential experience, as proposed in AP.

In the specification reported in columns (3) and (4) instead, we use tenure as time measure. In column (3) we treat tenure as exogenous. However, tenure depends on quit and layoff decisions, and therefore may be correlated with characteristics of workers and job matches. These same characteristics are likely to be related to worker productivity, and how fast employers learn about worker productivity. Therefore, we report in column (4) the result from a specification where we use an Instrumental Variables approach to deal with the heterogeneity bias. We use the variation of tenure over a given job match, following Altonji and Shakotko (1987), along with potential experience as instruments for job tenure. Specifically, our instruments are the deviations of the job tenure variables around their means for the sample observations on a given job match. This variable is by construction uncorrelated with both the individual and job specific unobserved components

Results show that the coefficients on AFQT and AFQT interacted with experience or tenure are all positive and significant, suggesting that productivity may be partially observed to employers at the time of initial hire and that employers learn about workers' productivity over time as they acquire new information. Overall the IV estimates are very similar to the OLS estimates. The coefficient of 0.065 (0.018)

Table 3: The Effects of AFQT and Race on Log Wages over Experience and Tenure

Panel 1–Non College Graduates				
	Actual Experience		Job Tenure	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Standard. AFQT	0.051***	0.037**	0.054***	0.053***
	(0.012)	(0.013)	(0.010)	(0.011)
AFQT×Exper/120	0.036***	0.052***		
	(0.011)	(0.011)		
AFQT×Tenure/120			0.065***	0.067***
			(0.018)	(0.015)
Black	-0.046*	-0.055*	-0.127***	-0.098***
	(0.021)	(0.023)	(0.019)	(0.019)
Black×Exper/120	-0.042*	-0.042		
	(0.020)	(0.023)		
Black×Tenure/120			0.049	-0.042
			(0.037)	(0.025)
R^2	0.258	0.253	0.253	0.251
No. of obs.	247,140			

Panel 2–College Graduates				
	Actual Experience		Job Tenure	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Standard. AFQT	0.123***	0.119***	0.156***	0.133***
	(0.026)	(0.027)	(0.024)	(0.022)
AFQT×Exper/120	0.038	0.038		
	(0.024)	(0.025)		
AFQT×Tenure/120			-0.036	0.025
			(0.041)	(0.032)
Black	0.138	0.163**	0.104*	0.104*
	(0.047)	(0.051)	(0.046)	(0.043)
Black×Exper/120	-0.091*	-0.129**		
	(0.038)	(0.046)		
Black×Tenure/120			-0.155*	-0.159**
			(0.077)	(0.053)
R^2	0.268	0.257	0.262	0.261
No. of obs.	70,848			

Notes: the experience measure is actural work experience in months. All specifications control for year effects, urban residence, region of residence. Specifications with experience also control for a quadratic term in actural experience, and specifications with tenure control for a quadratic term in tenure. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

on AFQT-tenure interaction term in specification (3) is positive and significant, consistent with the prediction of the employer learning model.

Recall that if employer learning is asymmetric, learning takes place mostly on job tenure as outside firms have little information regarding a worker's productivity. The coefficient on the AFQT-tenure interaction is greater than the estimated coefficient on AFQT-experience interaction in specification (1), with a P-value of 0.071. The P-value of the same comparision is larger on the IV Specifications (0.208). Overall, this evidence indicates that employer learn over time about workers skills, that learning re-starts at the beginning of each new job and that the speed of learning is faster over job tenure, providing evidence in favor of asymmetric learning.

Turning to the analysis of racial differences, we find that at the time of initial entry into the labor market, and in contrast to the results from AP, black workers earn less than white workers with the same AFQT score in all of our specifications, supporting the hypothesis that employers have limited information about the productivity of new workers and statistically discriminate on the basis of race.[16] While employer learning makes wages more correlated with skill over time, we do not find a strong evidence that the wage gap (conditional on measured skill) decreases over time: the coefficients of race interacted with tenure or experience are either negative or insignificant.

In Panel 2 of Table 3 we report the corresponding regression results for college graduates. The coefficients on AFQT are large and statistically significant but the coefficients on AFQT interacted with time are insignificant and relatively small in all specifications. These results are robust when we use actual experience and job tenure as alternative time measure as in columns (1) and (3), and when we treat experience and tenure as endogenous as in columns (2) and (4). The time trend of the returns to AFQT shows that there are substantial returns to AFQT for college graduate workers immediately after they take a new job. A one standard deviation increase in AFQT is associated with between 11.9–15.6% increase in wages. Moreover, the returns to AFQT are hardly affected by experience or tenure. Following the interpretation by Arcidiacono et al. (2010), the estimated AFQT-time profiles suggest that employers have accurate information about the productivity of newly hired college graduate workers and learn very little additional information over time.

In contrast to non-college graduates, college-educated black workers earn a higher

[16] AP find little evidence on statistical discrimination on the basis of race and argue that statistical discrimination plays a relatively unimportant role in the racial wage gap. When we pool the education groups, we also find little evidence on racial statistical discrimination. Mansour (2012) confirms AP's finding, but his empirical results imply that the pattern might differ across occupations.

wage than their white counterparts when they start a job, but this black wage premium declines over time.[17] Arcidiacono et al. (2010) argue that information contained on the resumes of college graduates, such as grades, majors and the college attended, help college-educated workers directly reveal their productivity to their employers. Therefore in the market for college graduate workers, employers have less incentives to statistically discriminate against black workers because they can assess workers' productivity more accurately at the time of initial hire. One plausible explanation for the black wage premium among college graduates is that black college workers are more motivated and productive than their white counterparts. If the AFQT and other tests such as SAT are racially biased, then blacks will have higher productivity than whites conditional on the test scores.[18] The diminishing black wage premium over time among high-skilled workers indicates that black workers may still suffer from racial prejudice in opportunities for promotion or on-the-job training over their careers even if there is no statistical discrimination at hiring.

We conclude that employer learning process mainly occurs in the market for non-college graduate workers. Productivity is observed nearly perfectly for workers with a college degree at hiring, and thus little scope is left for employer learning.

4.2 Job transitions

Our empirical results in Table 3 reveal that whenever non-college black workers start a new job, employers pay them significantly less than their white counterparts conditional on their AFQT scores. We would expect that non-college black workers generally switch firms less frequently than do non-college white workers, in accordance with Empirical Implication 5. For comparison with the analysis in the previous subsection, we estimate the job change probabilities for the non-college sample and the college sample separately.

The results of the probit regressions are presented in Table 4. In column (1), we include the black dummy, years of schooling, and actual work experience in the probit regression for non-college graduates and find no racial difference in job change probability. However when AFQT score is included in the regression in column (2), we find that non-college black workers are less likely to change jobs compared to white workers with the same education, experience and AFQT scores. Using the

[17] The existence of a substantial black wage premium for college graduates is a robust feature of the U.S. labor market. Neal and Johnson (1996) find that the racial wage gap for males declines with the skill level, and a similar finding is also reported in Lang and Manove (2011).

[18] As argued by Arcidiacono (2005), affirmative action in the workplace may also account for the inital black wage premium. Black workers earn more because the number of blacks with a college of degree is small, yet employers value diversity in the workplace.

Table 4: Probit Estimates on Racial Difference in Job Change Probabilities

	Non-College Graduates		College Graduates	
	(1)	(2)	(3)	(4)
Black	-0.0006	-0.0056***	0.0069**	0.0026
	(0.0009)	(0.0010)	(0.0021)	(0.0025)
Education	-0.0006*	0.0009**	-0.0005	0.0000
	(0.0003)	(0.0003)	(0.0006)	(0.0007)
Exper/120	-0.0247***	-0.0250***	-0.0305***	-0.0310***
	(0.0014)	(0.0014)	(0.0029)	(0.0029)
AFQT		-0.0058***		-0.0049***
		(0.0006)		(0.0013)
No. of obs.	798,108	798,108	173,160	173,160

Notes: the experience measure is actural work experience in months. All specifications control for year effects, urban residence, region of residence. We report the marginal effect of each variable. The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

coefficients from the probit regression in column (2), we calculate the marginal effect of the black dummy on the job transition probability evaluated at sample means. We find that black non-college workers are 0.56% less likely to make a month-to-month job change compared to their white counterparts.

In the non-college market, black workers face statistical discrimination at each time of new hire; therefore, black workers tend to change jobs less frequently than do white workers to mitigate the effects of discrimination. The results in column (2) also indicates that workers with higher AFQT scores change jobs less frequently, as they may be able to find higher quality matches. The average AFQT score of black workers is approximately one standard deviation lower than that of white workers, thus they are more likely to change jobs. On the other hand, black workers have less incentive to change jobs to avoid racial prejudice. As a result, we find no race effect in column (1) when AFQT score is not controlled for. These results are consistent with Empirical Implication 4 and provide further evidence in favor of the asymmetric learning and statistical discrimination hypotheses.

The results for college graduates appear in columns (3) and (4) of Table 4. Black college workers change jobs more frequently than white workers based on the estimates in column (3), but it is primarily because black workers have lower AFQT scores. Black college workers' month-to-month job mobility rates are not statistically different from those of white college workers conditional on AFQT scores as shown in column (4), and the results are robust with additional controls in column (6). We also do not find evidence that college-educated black workers have lower quality matches than white workers. Consistent with our previous finding that employers have little incentive to discriminate against black college workers at hiring time as employers can directly observe key aspects of workers' productivity, we do not find black college graduates to change job less frequently to mitigate discrimination.

To summarize, we find that black non-college graduates change job less frequently than do white non-college graduates in support of the asymmetric learning hypothesis. Furthermore, we find that there are no statistically significant racial differences in job mobility patterns in the college market, confirming that employers have more accurate information about college-educated workers at the time of hiring, and therefore do not statistically discriminate against blacks.

5 Additional results

5.1 Non-Purely Asymmetric Learning

As a robustness check and to examine the possibility that employer learning is not purely asymmetric, We analyze how AFQT varies with experience and tenure when both are included in the regression model.[19] If learning is symmetric, then the coefficient on AFQT-tenure interaction should be zero because employer learning process takes place over general experience. If learning is purely asymmetric, outside firms are completely excluded from the learning process. We should only observed learning over tenure, thus AFQT-experience interaction should be zero. Otherwise if learning is not purely asymmetric, some new productivity information is revealed to outside firms but more information is available to current firms and both AFQT-experience and AFQT-tenure interactions should have non-zero coefficients. We specify the following wage regression that includes both experience and tenure interaction terms for non-college graduates.

$$\begin{aligned}\ln w_i = {} & \beta_0 + \beta_S S_i + \beta_{S,X}(S_i \times X_i) + \beta_{S,T}(S_i \times T_i) \\ & + \beta_{AFQT} AFQT_i + \beta_{AFQT,X}(AFQT_i \times X_i) + \beta_{AFQT,T}(AFQT_i \times T_i) \\ & + \beta_{Black} Black_i + \beta_{Black,X}(Black_i \times X_i) + \beta_{Black,T}(Black_i \times T_i) \\ & + \beta_\Omega \Omega_i + H(X_i) + \epsilon_i, \end{aligned} \tag{7}$$

[19] Using a sample of white males from 1979–2001 waves of NLSY79, Schönberg (2007) examines whether employer learning is symmetric or imperfectly asymmetric by analyzing how education and AFQT vary with experience and tenure when both are included in the wage regression.

Table 5: Testing for Imperfect Asymmetric Learning for Non-College Graduates

	(1)	(2)
Standard. AFQT	0.054***	0.046***
	(0.010)	(0.012)
AFQT×Tenure/120	0.065***	0.050*
	(0.018)	(0.022)
AFQT×Exper/120		0.018
		(0.013)
Black	-0.127***	-0.059**
	(0.019)	(0.021)
Black×Tenure/120	0.049	0.097*
	(0.037)	(0.043)
Black×Exper/120		-0.065**
		(0.023)
R^2	0.253	0.277
N	247,140	247,140

Note: see notes to Table 3

The main coefficients of interest are $\beta_{AFQT,X}$, the coefficient on AFQT-experience interaction term, and $\beta_{Black,X}$, the coefficient on *Black*-experience interaction term. Pure asymmetric learning predicts that both $\beta_{AFQT,X}$ and $\beta_{Black,X}$ should equal zero whereas non-purely asymmetric learning indicates non-zero coefficients on experience interaction terms.

We report the estimates of equation (7) in Table 5, column (2). We present in column (1) for ease of comparison the results from the OLS specification (column 3 of Table 3). When both tenure and experience interactions are included as regressors, the coefficient on the AFQT-Experience interaction term is not statistically different from zero while the coefficient on the AFQT-Tenure interaction remains significant, empirical evidence in favor of purely asymmetric learning. This suggests that outside firms have little access to new information about workers' productivity as measured by AFQT over time.

When both black-tenure and black-experience interaction terms are included in the wage equation (7), the initial black coefficient becomes smaller but remains statistically significant. The significantly positive coefficient on the black-tenure interaction indicates that current firm learns about black workers' productivity over time and rely less on the race information to infer their productivity. On the other hand, the significantly negative coefficient on black-experience interaction is consistent with outside firms not learning about black workers' true productivity over time. Black workers without a college degree appear to be discriminated against more on jobs that require more work experience. These results provide supporting evidence for the assumption of purely asymmetric learning in the non-college labor market.

5.2 Occupation and Industry

It is well documented that workers of different demographic types and different skills sort themselves into different sectors in the labor market Heckman and Sedlacek (1985). If black and white workers sort themselves into jobs that require different skill levels or sectors that pay different wages, observed wage differences may be due to factors different from those implied by the learning-based statistical discrimination model we used. One alternative explanation is that black workers are more likely to be hired into jobs and sectors that pay lower wages at the start of their career and to be trapped in such jobs. The initial job assignments and sector allocations could influence the menu of workers' career paths. What appears as evidence of statistical discrimination could be due to differential job sorting by black and white workers.[20]

To test the possibility that racial wage gap is driven by blacks and whites being sorted into jobs of different skill levels, we add initial occupation to the estimating equations (4) and (5) as an additional control, and repeat the empirical analysis separately for non-college graduates and college graduates.[21] The regression results are presented in columns (1) and (4) of Table 6. In the non-college market (Panel 1), we find evidence of asymmetric employer learning and statistical discrimination even after controlling for initial occupations black and white workers take. Wages become more correlated with AFQT over time, and employer learning is faster over job tenure than over actual experience. The *Black* coefficient is initially negative and significant and rises (but insignificantly) with tenure consistently across specifications, providing evidence that our main results can not be attributed to differences in occupation sorting of different racial groups. Including the initial occupation in the regressions also does not alter the results for college graduates presented in Panel 2. College-educated blacks earn an initial wage premium conditional on their AFQT.[22]

We also explore the role of sector allocation by examining the effect of initial industry on the observed racial wage gap.[23] We repeat the empirical analysis for the

[20]Racial differences in the initial job assignments and sector allocations could also be an outcome of discrimination, which will strengthen our results.

[21]We distinguish 7 occupations: professional workers; managers; sales workers; clerical workers; craftsman and operatives; agricultural labors; and service workers.

[22]The results shown in Table 6 provide evidence for race-based statistical discrimination within occupations. Mansour (2012) finds that there is substantial variation in the time path of black coefficients across occupations. Therefore, the extent of racial statistical discrimination may vary across occupations.

[23]We distinguish 12 industries: agriculture; mining; construction; manufacturing; transportation,

Table 6: The Effects of Race on Log Wages Controlling for Initial Occupation and Industry

Panel 1–Non College Graduates

	Actual Experience			Job Tenure		
	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	0.034**	0.042**	0.037**	0.033**	0.042***	0.036**
	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
AFQT×Exper/120	0.038***	0.036***	0.036***			
	(0.011)	(0.011)	(0.011)			
AFQT×Tenure/120				0.075***	0.068***	0.070***
				(0.020)	(0.020)	(0.020)
Black	-0.066**	-0.072**	-0.062*	-0.140***	-0.146***	-0.137***
	(0.025)	(0.024)	(0.024)	(0.023)	(0.023)	(0.022)
Black×Exper/120	-0.033	-0.036	-0.036			
	(0.021)	(0.020)	(0.020)			
Black×Tenure/120				0.067	0.057	0.065
				(0.041)	(0.041)	(0.040)
Initial Occupation	Yes	No	Yes	Yes	No	Yes
Initial Industry	No	Yes	Yes	No	Yes	Yes
R^2	0.276	0.286	0.294	0.271	0.281	0.290
No. of obs.	189,120	189,120	189,120	189,120	189,120	189,120

Panel 2–College Graduates

	Actual Experience			Job Tenure		
	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	0.091***	0.108***	0.086**	0.141***	0.150***	0.136***
	(0.027)	(0.026)	(0.026)	(0.025)	(0.025)	(0.024)
AFQT×Exper/120	0.040	0.042	0.041			
	(0.024)	(0.024)	(0.023)			
AFQT×Tenure/120				-0.076	-0.052	-0.074
				(0.041)	(0.043)	(0.041)
Black	0.133**	0.124**	0.111*	0.105*	0.101*	0.091*
	(0.048)	(0.047)	(0.045)	(0.047)	(0.047)	(0.045)
Black×Exper/120	-0.109**	-0.097*	-0.101*			
	(0.040)	(0.040)	(0.039)			
Black×Tenure/120				-0.221**	-0.201**	-0.217**
				(0.075)	(0.077)	(0.073)
Initial Occupation	Yes	No	Yes	Yes	No	Yes
Initial Industry	No	Yes	Yes	No	Yes	Yes
R^2	0.318	0.317	0.343	0.310	0.305	0.333
No. of obs.	65,376	65,376	65,376	65,376	65,376	65,376

Note: see notes to Table 3

two educational groups of interest with initial industry included as a control (columns 2 and 5 in Table 6). The results closely resemble those without the inclusion of initial industry in Table 3. Finally we control for initial occupation and industry simultaneously (columns 3 and 6 of the same table). The main results are not affected. Therefore, the racial wage gap can not be explained by the variations in initial occupations or industries that members from different racial groups work in.[24]

6 Conclusion

In this paper, we combine elements of both employer learning and statistical discrimination theories formulating a framework that nests both symmetric and asymmetric employer learning, and examine whether employers statistically discriminate against black workers at time of hiring under each scenario.

Our estimation results show that non-college graduates and college graduates are associated with different patterns of employer learning. At the time of initial hire, employers have to rely on some easily observable characteristics to estimate the productivity of non-college graduates, and they gradually update their expectations as they acquire more information. The time paths of racial wage gap in the non-college market indicate that employers use race as information to infer workers' productivity and black workers are statistically discriminated. We find that learning correlates more with tenure than with experience, and that blacks, who are initially statistically discriminated, are less likely to change jobs, supporting the hypothesis that learning is asymmetric in the non-college workers labor market. Among college graduates, we do not find evidence that african-american workers are statistically discriminated.

Many statistical discrimination models build on the assumption that the signal of productivity employers receive from black workers is less reliable than that from white workers at the time of initial hire[25]. Pinkston (2009) applies the framework of employer learning to test this hypothesis, and his estimation results provide evidence supporting this view.[26] Our learning-based racial statistical discrimination model assumes that the signals sent by workers from different racial groups are equally

communication, and utilities; wholesale and retail trade; finance, insurance, and real estate; business and repair services; personnel services; entertainment and recreation services; professional and related services; and public administration.

[24]Table 76presents the OLS estimates of the wage regressions. Results from the IV estimates treating actual experience or job tenure as endogenous are similar and available upon request.

[25]See e.g., Aigner and Cain (1977), Lundberg and Startz (1983), Cornell and Welch (1996), Lang (1986)

[26]Flabbi et al. (2016) provide empirical support to the hypothesis that signal quality differs by gender.

informative. An interesting topic for future research is to relax the assumption of equally informative signals from different racial groups and to investigate its effect on employer learning and statistical discrimination.

Finally, we note that our paper, as all of the related literature, is not designed to measure to what extent the persistent racial wage differences are due to statistical versus "taste-based" discrimination (in the sense of Becker (1971))[27]. Disentangling the different source of group inequality remains an important topic of future research.

A Appendix: derivation of expected productivities

Before hiring, the first signal s_0 is observed. The conditional distribution of productivity given the signal is normal, with

$$\begin{aligned} E(q|s_0) &= (1-\alpha)\mu + \alpha s_0, \text{ with } \alpha = \frac{\sigma^2}{\sigma^2+\sigma_\epsilon^2} \\ Var(q|s_1) &= \frac{\sigma^2\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} = \alpha\sigma_\epsilon^2 \end{aligned}$$

In period 1, the firm has another signal available, s_1 normality of the conditional distribution is preserved with the following moments following the same rule:

$$\begin{aligned} E(q|s_0,s_1) &= (1-k_1)E(q|s_0) + k_1 s_1, \text{ with } k_1 = \frac{\alpha\sigma_\epsilon^2}{\alpha\sigma_\epsilon^2+\sigma_\epsilon^2} = \frac{\alpha}{\alpha+1} \\ &= (1-k_1)(1-\alpha)\mu + (1-k_1)\alpha s_0 + k_1 s_1 \\ &= \frac{1}{\alpha+1}(1-\alpha)\mu + \frac{1}{\alpha+1}\alpha(s_0+s_1) \\ &= \frac{1-\alpha}{1+\alpha}\mu + \frac{2\alpha}{1+\alpha}\frac{\sum s_i}{2} = \\ &= \left(1-\frac{2\alpha}{1+\alpha}\right)\mu + \frac{2\alpha}{1+\alpha}\frac{\sum s_i}{2} \\ &= (1-\alpha_1)\mu + \alpha_1\overline{s} \text{ with } \alpha_1 = \frac{2\alpha}{1+\alpha} \\ Var(q|s_0,s_1) &= \frac{\alpha}{1+\alpha}\sigma_\epsilon^2 \end{aligned} \tag{8}$$

In period 2, another signal becomes available, s_2, we follow the same steps:

[27] See Neal and Johnson (1996) and surveys by Altonji and Blank (1999) and Lang and Lehmann (2012), among others

$$
\begin{aligned}
E(q|s_0, s_1, s_2) &= (1-k_2)E(q|s_0, s_1) + k_2 s_2, \text{ with } k_2 = \frac{\frac{\alpha}{1+\alpha}\sigma_\epsilon^2}{\frac{\alpha}{1+\alpha}\sigma_\epsilon^2 + \sigma_\epsilon^2} = \frac{\alpha}{2\alpha+1} \\
&= (1-k_2)(1-\alpha_1)\mu + (1-k_2)\alpha_1 \frac{s_1+s_2}{2} + k_2 s_2 \\
&= \frac{1}{2\alpha+1}(1-\alpha)\mu + \frac{3\alpha}{2\alpha+1}\frac{s_1+s_2+s_3}{3} \\
&= (1-\alpha_2)\mu + \alpha_2 \overline{s} \text{ with } \alpha_2 = \frac{3\alpha}{1+2\alpha} \\
Var(q|s_0, s_1, s_2) &= \frac{\alpha}{1+2\alpha}\sigma_\epsilon^2
\end{aligned}
$$

Hence, by induction,

$$
\begin{aligned}
E(q|s_0, s_1 ... s_T) &= (1-\alpha_T)\mu + \alpha_T \overline{s} \text{ with } \alpha_T = \frac{(T+1)\alpha}{1+T\alpha} \\
Var(q|s_0, s_1 ... s_T) &= \frac{\alpha}{1+T\alpha}\sigma_\epsilon^2
\end{aligned}
$$

Consider an employer hiring someone employed by another firm with tenure $T-1$. Assume asymmetric learning, that is at time T the new employer observes a signal ν that carries less precise information than the information available to the incumbent employer had she observed an extra signal s_T:

$$
\begin{aligned}
E(q|v) &= (1-\alpha^o)\mu + \alpha^o s_0, \text{ with } \alpha^o = \frac{\sigma^2}{\sigma^2 + \sigma_v^2} < \alpha_T \\
Var(q|v) &= \frac{\sigma^2 \sigma_v^2}{\sigma^2 + \sigma_\nu^2} = \alpha^o \sigma_\nu^2
\end{aligned}
$$

It is also reasonable to assume that $\sigma_\nu \leq \sigma_\epsilon$ that is the first piece of information available to the new employer is at least as good as the first piece of information available to the first employer (s_0).

In the first period after the new hire we nave

$$
\begin{aligned}
E(q|v, s_1) &= (1-\alpha_1^o)E(q|\nu) + \alpha_1^o s_1, \text{ with } \alpha_1^o = \frac{\alpha^o \sigma_\nu^2}{\alpha^o \sigma_\nu^2 + \sigma_\epsilon^2} \\
&= (1-\alpha_1^o)\left((1-\alpha^o)\mu + \alpha^o \nu\right) + \alpha_1^o s_1 \\
Var(q|\nu, s_1) &= \alpha_1^o \sigma_\epsilon^2
\end{aligned}
$$

In comparing the information available to the new employer after the first period and the information available to an employer with a worker with the same experience $T+1$, but tenure equal to $T+1$, note that because $\alpha^o < \alpha^T$ and $\sigma_\nu < \sigma_\epsilon$ then

$\alpha_1^o < k_{T+1}$, where k_{T+1} is the weight placed on the last signal received (see 8). Therefore, the weight placed on the unconditional mean μ is greater for workers that stay with the same employer and have the same experience but longer tenure.

The same logic carries over when new signals are introduced

$$
\begin{aligned}
E(q|v, s_1, s_2) &= (1 - k_2^o)E(q|v, s_1) + k_2^o s_2 \text{ with } k_2^o = \frac{\alpha_1^o \sigma_\epsilon^2}{\alpha_1 \sigma_\epsilon^2 + \sigma_\epsilon^2} = \frac{\alpha_1^o}{\alpha_1^o + 1} \\
&= (1 - k_2^o)(1 - \alpha_1^o)\left((1 - \alpha^o)\mu + \alpha^o \nu\right) + (1 - k_2^o)\alpha_1^o s_1 + k_2^o s_2 \\
&= \left(1 - \frac{2\alpha_1^o}{1 + \alpha_1^o}\right)\left((1 - \alpha^o)\mu + \alpha^o \nu\right) + \frac{2\alpha_1^o}{1 + \alpha_1^o} \frac{\sum_{i=2,3} s_i}{2} \\
&= (1 - \alpha_2^o)\left((1 - \alpha^o)\mu + \alpha^o \nu\right) + \alpha_2^o \overline{s} \text{ with } \alpha_2^o = \frac{2\alpha_1^o}{1 + \alpha_1^o} \\
Var(q|v, s_1, s_2) &= \alpha_2^o \sigma_\epsilon^2
\end{aligned}
$$

And then again adding additional signals one can derive as above:

$$
\begin{aligned}
E(q|\nu, s_1 ... s_T) &= (1 - \alpha_T^o)\left((1 - \alpha^o)\mu + \alpha^o \nu\right) + \alpha_T \overline{s} \text{ with } \alpha_T^o = \frac{T^o \alpha}{1 + (T^o - 1)\alpha} \\
Var(q|\nu, s_1 ... s_T) &= \frac{\alpha}{1 + T\alpha} \sigma_\epsilon^2
\end{aligned}
$$

References

Aigner, Dennis J and Glen G Cain, "Statistical theories of discrimination in labor markets," *ILR Review*, 1977, *30* (2), 175–187. 2, 25

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 2001, *116* (1), 313–350. 2

Altonji, Joseph G and Rebecca M Blank, "Race and gender in the labor market," *Handbook of labor economics*, 1999, *3*, 3143–3259. 26

Altonji, Joseph G. and Robert A. Shakotko, "Do Wages Rise with Job Seniority?," *The Review of Economic Studies*, 1987, *54* (3), 437–459. 16

Arcidiacono, Peter, "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524. 19

— , **Patrick Bayer, and Aurel Hizmo**, "Beyond signaling and human capital: Education and the revelation of ability," *American Economic Journal: Applied Economics*, 2010, *2* (4), 76–104. 2, 4, 11, 15, 18, 19

Arrow, Kenneth, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton, N.J.: Princeton University Press, 1973, pp. 3–33. 2

Bauer, Thomas K and John P Haisken-DeNew, "Employer learning and the returns to schooling," *Labour Economics*, 2001, *8* (2), 161–180. 4

Becker, Gary, *The economics of discrimination*, Chicago: The University of Chicago Press (first ed. 1957), 1971. 2, 26

Bernhardt, Dan, "Strategic promotion and compensation," *The Review of Economic Studies*, 1995, *62* (2), 315–339. 2

Coate, Stephen and Glenn C. Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *The American Economic Review*, December 1993, *83* (5), 1220–1240. 2

Cornell, Bradford and Ivo Welch, "Culture, information, and screening discrimination," *Journal of Political Economy*, 1996, pp. 542–571. 25

Eaton, Morris L., *Multivariate Statistics : A Vector Space Approach* Wiley series in probability and mathematical statistics, New York : Wiley, 1983. 5

Fang, Hanming and Andrea Moro, "Theories of Statistical Discrimination and Affirmative Action: A Survey," in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics*, Vol. 1A, The Netherlands: North Holland, 2011, chapter V, pp. 133–200. 2

Farber, Henry S. and Robert Gibbons, "Learning and Wage Dynamics," *The Quarterly Journal of Economics*, 1996, *111* (4), 1007–1047. 2

Flabbi, Luca, Mario Macis, Andrea Moro, and Fabiano Schivardi, "Do Female Executives Make a Difference? The Impact of Female Leadership on Gender Gaps and Firm Performance," Working Paper 22877, National Bureau of Economic Research December 2016. 25

Gibbons, Robert and Lawrence F Katz, "Layoffs and lemons," *Journal of labor Economics*, 1991, *9* (4), 351–380. 4

Golan, Limor, "Counteroffers and efficiency in labor markets with asymmetric information," *Journal of Labor Economics*, 2005, *23* (2), 373–393. 2

Greenwald, Bruce C., "Adverse selection in the labour market," *The Review of Economic Studies*, 1986, *53* (3), 325–347. 2

Heckman, James J and Guilherme Sedlacek, "Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market," *Journal of Political Economy*, 1985, *93* (6), 1077–1125. 23

Kahn, Lisa B, "Asymmetric Information Between Employers," *American Economic Journal: Applied Economics*, 2013, *5* (4), 165–205. 4

Lang, Kevin, "A language theory of discrimination," *The Quarterly Journal of Economics*, 1986, pp. 363–382. 25

_ **and Jee-Yeon K Lehmann**, "Racial discrimination in the labor market: Theory and empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006. 2, 26

_ **and Michael Manove**, "Education and labor market discrimination," *American Economic Review*, 2011, *101* (4), 1467–96. 19

Lange, Fabian, "The Speed of Employer Learning," *Journal of Labor Economics*, 2007, *25* (1), 1–35. 2

Lesner, Rune V, "Testing for Statistical Discrimination based on Gender," *Labour*, 2018, *32* (2), 141–181. 2

Lundberg, Shelly J and Richard Startz, "Private discrimination and social intervention in competitive labor market," *The American Economic Review*, 1983, *73* (3), 340–347. 25

Mansour, Hani, "Does employer learning vary by occupation?," *Journal of Labor Economics*, 2012, *30* (2), 415–444. 2, 18

Neal, Derek A and William R Johnson, "The role of premarket factors in black-white wage differences," *Journal of political Economy*, 1996, *104* (5), 869–895. 19, 26

Phelps, Edmund S, "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 1972, *62* (4), 659–661. 1, 3, 5

Pinkston, Joshua C., "A Model of Asymmetric Employer Learning with Testable Implications," *The Review of Economic Studies*, 2009, *76* (1), 367–394. 4, 7, 25

Schönberg, Uta, "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 2007, *25* (4), 651–691. 4, 21

Waldman, Michael, "Job assignments, signalling, and efficiency," *The RAND Journal of Economics*, 1984, *15* (2), 255–267. 2