Testing for Asymmetric Employer Learning and Statistical Discrimination

Suqin Ge, Andrea Moro, and Beibei Zhu*

June 25, 2018

Abstract

We test the implications of a statistical discrimination model in which firms learn about worker productivity over time and may use race to infer worker productivity. We allow for asymmetric employer learning, that is, we assume that 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. For college graduates, employers directly observe most of the productivity of potential employees at hiring and learn very little over time.

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 imperfectly informed employers use group average demographic characteristics as proxies of unobserved worker characteristics, such as their productivity and qualification.[1] At the time of hiring, the productivity of labor

*Ge: Department of Economics, Virginia Tech. Moro : Department of Economics, Vanderbilt University. Zhu: Amazon. This study was prepared with 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 or the Department of Labor. 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 some unexplained exogenous difference between groups of workers, combined with

force participants is difficult to observe directly, thus employers use easily observable characteristics such as race or gender to infer the unobservable productivity and statistically discriminate among their potential employees. In an influential paper, Altonji and Pierret (2001) combined 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 on the true productivity of individual worker. 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 (e.g., Lange (2007), Arcidiacono et al. (2010), and Mansour (2012)), but all 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 studies 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" in the literature.[3] The nature of employer learning may have great influence on the effects of statistical discrimination. The discriminated group suffers a wage loss, and learning has to restart at each job turnover if the new employer learns nothing from a worker's previous jobs, that is, if employer learning is purely asymmetric. On the contrary, symmetric learning implies a continuous learning process regardless of job turnovers. Our study develoops a testable model that nests both symmetric and asymmetric learning hypotheses and providing new tests for statistical discrimination based on race.

employers' imperfect information about workers' productivity. This literature (see also Aigner and Cain (1977)), is largely agnostic on where the initial group differences come from. They may result from either differences in employer perceptions or other factors, such as differences in school quality. The other branch of this theory, originated from Arrow (1973) and analyzed 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 standard 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 (such as race, gender, immigration status) arises because employers perceive the average productivity differences between groups. We extend this model and allow the employer learning process to be asymmetric, that is, outside firms have less information about worker productivity that current firms have learned.

We show that asymmetric learning implies different predictions about how wages evolve with 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. We show that this implication is the opposite when learning is symmetric. In addition, when employers statistically discriminate against minorities, initially black wages (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 (black) worker changes jobs. Hence we should expect fewer job transitions among black workers, conditional on their skills. This is not the case if learning is symmetric, because new employers inherit the learning that occurred from 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 statistically discriminate against black workers, and learn asymmetrically about their skills. 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, and the wage penalty decreases over time, consistently with statistical discrimination. In addition, and confirming the asymmetric learning hypothesis, we find higher mobility of non-college educated black workers conditional on AFQT.

Results are different for college educated workers. We find no evidence of either statistical discrimination, nor that learning is asymmetric for this class of 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, supporting the main findings reported in Arcidiacono et al. (2010).

In addition to the papers we already cited that belong the employer learning with 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 is, looking 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 under different scenarios. Section 3 gives an overview of the data and compares the results using the specification in AP using versions our sample. The main estimation results are reported in section 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. Workers have ability q which is distributed normally 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 environment, X includes group identity such as race or gender. Ability, productivity and wages are expressed in logarithms to guarantee that all of these variables are positive in levels. We abstract at this stage from treating the evolution of workers' human capital from increasing experience (which can be included in vector X, so that average productivity can be interpreted as conditional on experience as well).

Employers observe a signal of ability s_1 from an unemployed worker, and then, after hiring a worker, in each period t they observe from her additional signals $s_t = q + \epsilon_t$, 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 up to the value of workers' expected productivity[4].

It is helpful to start the analysis by exploring the effect of employer learning about workers' productivity. In each period, employers' compute the worker's expected productivity given the signals observed. New workers are offered expected productivity $E(q|s_1)$. Standard properties of the bivariate normal distribution[5] imply that

$$E(q|s_1) = (1-\alpha)\mu(X) + \alpha s_1,$$
$$\alpha = \frac{\sigma^2}{\sigma_\epsilon^2 + \sigma^2}.$$

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;

[4]As we discuss below, this assumption is not crucial

[5]See Eaton (1983)

when the signal is pure noise ($\sigma_\epsilon = \infty$), expected ability is equal to the population average. With a partially informative signal, 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 $\alpha\sigma_\epsilon^2$. Denote the corresponding cumulative distributions with $\Phi_g(q|s_1)$.

As the worker increases her tenure with a firm, the 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 the expected productivity is:

$$E(q|s_1, ..., s_T) = (1 - \alpha_T)\mu(X) + \alpha_T \left(\frac{\sum_t s_t}{T}\right) \tag{1}$$

where:

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

Note that $\alpha_1 = \alpha$, and that the weight α_T placed on the signal average is increasing in T and converges to 1. As tenure increases, the worker's expected productivity gets closer to her true productivity.

We assume in what follows that labor market competition ensures that wages are equal to the expected productivity given the vector of signals. Consider an econometrician observing wages and a one-time signal of skill r not observed by the employer, such as the Armed Forces Qualification Test (AFQT), such that $r = q+\epsilon_r$, $\epsilon_r \sim N(0, \sigma_r^2)$. Then,

$$cov(r, E(q|s_1...s_T)) = cov\left(q + \epsilon_r, (1 - \alpha_T)\mu(X) + \alpha_T \left(\frac{\sum_t (q + \epsilon_t)}{T}\right)\right) = \alpha_T Var(q)$$

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 [6].

Consider now the implications of the model for statistical discrimination. Assume there are two groups of workers that belong to different but recognizable groups: a

[6] 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.

minority (M) and a dominant (D) group. Assume that $\mu(M) < \mu(D)$ and that employers use race for labor market decisions. 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 2. 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.

Consider now workers that are hired after being separated from their employers. Workers draw a new match-specific productivity q^o from a normal distribution with mean q, and a new signal for the new employer. 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. Crucially, we also assume that the information available is worse than the information available to the current employer, which implies

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

where T is the number of years the worker has been with the current employer. We believe this assumption to be realistic because a worker's curriculum, job interviews, aptitude tests cannot substitute from day-to-day interactions over the workers' tenure. After being hired, employers' expectations evolve in the same way they do for new hires, that is, employers extract a signal every period, and revise their posterior expectations using Bayes' rule, that is

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

where

$$\alpha_T^o = \frac{T\alpha^o}{1 + (T-1)\alpha^o}$$

Compare for the sake of example, two workers, Mary and John, with the same experience T. Mary has been always with the same employer, whereas John has worked for two employers without labor market discontinuity, and has current tenure $S < T$. Then note that for Mary's employer the weight on Mary's signal stream is (2),

$$\alpha_T = \frac{T\alpha}{1 + (T-1)\alpha} \tag{4}$$

whereas for John's employer

$$\alpha_M^o = \frac{S\alpha^o}{1 + (S-1)\alpha^o}$$

with $\alpha_S^o < \alpha_T$ because from assumption (3) we have $\alpha^o < \alpha_{T-S}$: at the time when John changes employers, the new employers starts off trusting the signal less than Mary's employer. This implies that wages of workers with discontinuous work histories covary with the econometricians' signals less than workers with continuous work histories, leading to the following implication.

Empirical Implication 3. Consider two regressions that include the interactions of either tenure with AFQT, or experience with AFQT. The coefficient of the interaction of tenure with AFQT is positive and larger than the coefficient of the interaction of experience with AFQT. In a wage regression where tenure is controlled for, the coefficient on the interaction of experience with AFQT is zero.

Finally, we want to look at how incomplete information affects workers' job mobility. 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 $\mu(M) < \mu(Q)$. 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 $\sigma^2(X)$.[7] The new firm observes a noisy signal of q', $q' + \epsilon$, where as before $\epsilon \sim N(0, \sigma_\epsilon^2)$. 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

[7]We cannot simply assume that the worker retains her productivity q and draws a new signal with the new firm because, conditional on acceptance, the new firm is never willing to pay a wage that the worker will accept. Consider for example a worker with productivity q, expecting to earn q forever with her current employer. She will change jobs only if the expected stream of wages is at least the same in exoected present value. Hence, she will accept only if the wage offer is above q. But this implies that the expected productivity, conditional on acceptance, from the employers' point of view, is lower than the offered wage. A worker accepting an offer w has expected productivity equal to $\int_{-\infty}^{w} qf(q)dq < w$, hence a firm is never willing to make an offer. A worker that is paid by her current employer less than productivity q instead, may accept an offer below q, which would be profitable for the new employer. However, what is the expected profit for the new employer? From acceptance the employer can only infer that the worker was paid less before

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 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 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.[8] Hence, workers from group D will be more likely to change jobs. We can state the following

Empirical Implication 4. If group M is statistically discriminated against, conditional on AFQT 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, and illustrate the implication about sign and magnitudes of its regression coefficients. The main departure of our specification relative to previous literature is that we consider two different time measures. 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 measures. The corresponding log wage equations we

[8] 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

estimate are as follows:[9]

$$\ln w_i = \beta_0^X + \beta_S^X S_i + \beta_{S,X}^X (S_i \times X_i) + \beta_{AFQT}^X AFQT_i + \beta_{AFQT,X}^X (AFQT_i \times X_i) \\ + \beta_{Black}^X Black_i + \beta_{Black,X}^X (Black_i \times X_i) + \beta_{\Omega}^X \Omega_i + H\left(X_i\right) + \epsilon_i^X, \tag{5}$$

$$\ln w_i = \beta_0^T + \beta_S^T S_i + \beta_{S,T}^T (S_i \times T_i) + \beta_{AFQT}^T AFQT_i + \beta_{AFQT,T}^T (AFQT_i \times T_i) \\ + \beta_{Black}^T Black_i + \beta_{Black,T}^T (Black_i \times T_i) + \beta_{\Omega}^T \Omega_i + H\left(X_i\right) + \epsilon_i^T. \tag{6}$$

where w_i denotes the hourly wage of individual i, S_i measures years of schooling, $AFQT_i$ denotes his AFQT score, $Black_i$ is a dummy variable on race, and Ω_i 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 variable t_i measures time, and $H\left(t_i\right)$ is a polynomial in time. Time is measured in months in our sample, and we divide the interaction of any variable with time measure t by 120 so the coefficients on interaction terms measure the change in wage during a ten-year period.

We exploit now the empirical predictions from Section 2 to derive implications about the signs and magnitudes of some coefficient estimates from these equations. These implications are summarized by Table 1. Panel (A) reports the implications assuming symmetric learning, and Panel (B) reports implications under asymmetric learning.

We focus first on employer learning and its implication on the coefficients on AFQT and its interaction with tenure and experience. The coefficient β_{AFQT}^X measures the correlation of wages with skills for workers with zero experience. This coefficient should be positive if the first employer has any information about skills available. When learning is symmetric, β_{AFQT}^T should be higher than β_{AFQT}^X because there are workers with zero tenure with some experience, which implies that their employers had some opportunities to learn about their productivity (row 1). Under asymmetric learning instead, employers of workers with zero tenure have the same information employers of workers with zero experience have: there was no opportunity to learn from work histories, therefore the coefficients should be the same (row 5).

Following the interpretation of AP and our theoretical model, if employers have limited information of the productivity of labor market entrants, wages become more

[9]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.

Table 1: Predictions of the Employer Learning and Statistical Discrimination Model

	(A) Symmetric Learning	
(1)	$0 < \beta^X_{AFQT} < \beta^T_{AFQT}$	
(2)	$0 < \beta^T_{AFQT,T} < \beta^X_{AFQT,X}$	
	With Statistical Discrimination	Without Statistical Discrimination
(3)	$\beta^X_{Black} < \beta^T_{Black} < 0$	$\beta^T_{Black} < \beta^X_{Black} = 0$
(4)	$0 < \beta^T_{Black,T} < \beta^X_{Black,X}$	$\beta^X_{Black,X} < \beta^T_{Black,T} < 0$
	(B) Asymmetric Learning	
(5)	$0 < \beta^X_{AFQT} = \beta^T_{AFQT}$	
(6)	$0 < \beta^X_{AFQT,X} < \beta^T_{AFQT,T}$	
	With Statistical Discrimination	Without Statistical Discrimination
(7)	$\beta^X_{Black} = \beta^T_{Black} < 0$	$\beta^X_{Black} = \beta^T_{Black} = 0$
(8)	$0 < \beta^X_{Black,X} < \beta^T_{Black,T}$	$\beta^T_{Black,T} < \beta^X_{Black,X} < 0$

strongly correlated with AFQT as experience accumulates. This means that in equation (5) $\beta^X_{AFQT,X}$ should be positive. Our theory has further implications on the regression coefficients in equations (5) and (6). If learning is symmetric, using tenure instead of experience (equation 6) should result in a smaller coefficient because workers with with a low tenure may 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 (row 2). The opposite happens when learning is asymmetric, or based on tenure: some workers with high experience have low tenure (hence lower learning opportunities for their employer), which reduces the correspondig coefficient of a regression using experience (row 6).

We turn now to the implications on the coefficients on race, which depend on whether or not employers statistically discriminate. When employers statistically discriminate, the coefficient on the black dummy should be negative, with $\beta^X_{Black} < \beta^T_{Black} < 0$ when learning is symmetric, because some workers with zero tenure have positive experience, which improves their wages (row 3). If learning is asymmetric, then $\beta^X_{Black} = \beta^T_{Black} < 0$ because workers with zero tenure are in the same informational situation as workers with zero experience (row 7). If employers do not use race as information, then $\beta^X_{Black} = 0$ and the sign of β^T_{Black} depends on relative race pre-market skills. In our data blacks have lower pre-market skills. If employers do not use race as information, but use signal from workers, then over time wages will become more correlated with skills and because zero-tenure workers include workers with positive experience, we have $\beta^T_{Black} < 0$ when learning is symmetric and $\beta^T_{Black} = 0$ when it is asymmetric.

Finally, in row (4) and (8) we report the implications regarding the coefficients on the interaction of time with race. Over time, wages of black and white conditional on skills will converge, therefore these coefficients are positive under statistical discrimination and negative when employers do or cannot use race. With symmetric learning, the speed of employer learning is faster over experience than over tenure if the speed of employer learning declines as the worker's career progresses (as it is in the case under our distributional assumptions). Hence, the coefficients on experience interactions should be larger (in absolute values) than the coefficients on tenure interactions. However, in our empirical specifications in equations (5) and (6), the effects of AFQT and race on log wages are assumed to vary linearly with time. This may impose the coefficients on experience interactions in (5) to be equal to those on tenure interactions (6). Combining the implications from the theory and the linear specifications, we have $0 < \beta^T_{AFQT,T} \leq \beta^X_{AFQT,X}$, $0 < \beta^T_{Black,T} \leq \beta^X_{Black,X}$ with discrimination, and $0 > \beta^T_{Black,T} \geq \beta^X_{Black,X}$ without discrimination.

On the other hand, if employer learning is purely asymmetric, the initial employer and all subsequent employers have the same information regarding a worker's productivity. Therefore β^X_{AFQT} should equal β^T_{AFQT}, and β^X_{Black} should equal β^T_{Black}. Moreover, the speed of employer learning is faster over tenure than over experience because employer learning occurs over tenure, but it is interrupted over work experience after any job-to-job transition. As a result, the coefficients on tenure interaction terms should be larger (in absolute values) than the coefficients on experience interaction terms. This means that $\beta^T_{AFQT,T} \geq \beta^X_{AFQT,X} > 0$, $\beta^T_{Black,T} \geq \beta^X_{Black,X} > 0$ if there exists race-based discrimination, and $\beta^T_{Black,T} \leq \beta^X_{Black,X} < 0$ if there is no discrimination. Table 2 summarizes the predictions on the regression coefficients from our model depending on the nature of employer learning and the existence of race-based statistical discrimination.

3 Data

The empirical analysis is based on the 2008 release of NLSY79. The NLSY79 is 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 are then interviewed on a biennial basis. The NLSY 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 rich 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 as in Altonji and Pierret (2001). 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 and at least three weeks in the month, 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 a given employer. We link all the jobs across different 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 when he is back in nonemployment. 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.[10] 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 mor than $100 per hour are excluded from the analysis. We use the AFQT as our

[10] In Altonji and Pierret (2001), 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.

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.[11] The total remaining sample consists of 2,595 whites and 1,136 blacks with 320,124 monthly observations.

Table 2 presents the summary statistics for the main variables in our sample by race and education level. We consider 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).[12] 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 consider blacks and whites of the same 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 they use the information on average AFQT differences between blacks and whites. In the next section, we carry out the empirical analysis to examine this issue in detail.

3.1 Comparison of Altonji and Pierret (2001)'s results with our specification

In this section, we begin our empirical analysis with a replication of the results reported in AP using our sample selection criteria. We estimate a log earning equation in the form of (5) and present the results in Table 3. We report for convenienence in Column (1) the results from AP, Table 1, Panel 1, Column 4. The specification in column (2) using data from the same time period, interview years 1979–1992, but with several differences in sample construction. First, we use monthly data instead of annual data. Secondly, the construction of potential experience is slightly different. We measure potential experience as time since first left school instead of age minus years of schooling minus six, the experience measure in AP. Despite the slight differences, the main qualitative results are still present.

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

[12] 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 empirical results are similar. As both high school dropouts and workers with some college education but without a college degree behave similar to high school graduates, we bundle them into a sample of workers with no college degree.

Table 2: Summary Statistics by Race

	Whites			Blacks		
	All	<College	≥College	All	<College	≥College
AFQT	0.502	0.202	1.344	-0.568	-0.726	0.486
	(0.957)	(0.885)	(0.569)	(0.798)	(0.648)	(0.894)
Education (yrs)	13.35	12.15	16.74	12.70	12.12	16.60
	(2.39)	(1.31)	(1.19)	(2.01)	(1.36)	(1.06)
Hourly wage	12.94	11.11	18.08	10.19	9.24	16.56
	(8.21)	(5.90)	(11.08)	(6.19)	(4.99)	(9.03)
Experience:						
Potential	132.29	135.88	122.25	146.35	148.46	132.21
	(85.03)	(86.68)	(79.37)	(86.21)	(86.89)	(80.13)
Actual	111.29	112.67	107.43	111.84	111.74	112.50
	(76.49)	(77.93)	(72.18)	(73.53)	(73.89)	(71.12)
Job tenure	47.09	46.11	49.83	41.06	40.16	47.11
	(48.59)	(48.44)	(48.89)	(43.58)	(42.75)	(48.31)
Individuals	2,593	1,905	688	1,136	989	147
Observations	225,708	166,392	59,316	94,416	82,116	12,300

Standard deviations in parenthesis. Education is measured in years, real hourly wages in ?? dollars, experience in months, potential experience is months since left school.

Table 3: The Effects of Schooling, AFQT and Race on Log Wages

	(1)	(2)	(3)	(4)
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*
	(0.06)	(0.031)	(0.015)	(0.026)
R^2	0.287	0.273	0.346	0.322
Sample	AP (1979–1992)	AP sample (1979–1992)	Full sample (1979–2008)	Full sample Exp. ≤168m
No. of Observations	21,058	177,288	317,988	212,640

Notes: 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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The coefficient on education is positive and significant initially and falls over time, consistent with statistical discrimination on the basis of education. The 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 using the longer sample, 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. In our empirical analysis below, we follow the literature and assume the effects of AFQT and *Black* on log wages to vary linearly with time to keep the interpretation of these coefficients simple.

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.[13] We use these two samples from NLSY 1979–2008 to test the main predictions of our learning-based statistical discrimination model.

Our empirical analysis has two main focal points. First, to distinguish between

[13] Arcidiacono, Bayer and Hizmo (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 presented in Table 4.

symmetric and asymmetric employer learning, we examine the initial coefficients on AFQT and race as well as their interaction terms with time when experience and tenure are used as two separate time measures in log wage equations (5) and (6). According to Table 1, different initial coefficients and larger (in absolute value) coefficients on experience interaction terms are supportive for symmetric learning whereas asymmetric learning implies similar initial coefficients and larger (in absolute value) coefficients on tenure interaction terms because the speed of learning is faster over tenure when learning is asymmetric. Secondly, 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 4 we report estimates of wage regressions using the non-college graduate sample. If employer learning is symmetric, learning takes place over general work experience. Specification (1) estimates equation (5) 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 a 2SLS regressions where actual experience is instrumented with potential experience, as proposed in AP.

Results show that he coefficients on AFQT and AFQT interacted with experience are both 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. In addition, at the time of initial entry into the labor market, black workers earn wages approximately 4.7 percent less than those received by their white counterparts with the same AFQT score, providing some evidence for statistical discrimination on the basis of race in the market for non-college graduates.

Results from the IV regressions reported in column (2) are basically consistent with those in column (1). The IV estimate of the coefficient on AFQT is smaller

Table 4: The Effects of AFQT and Race on Log Wages over Experience and Tenure

Panel 1–Non College Graduates

	Actual Experience		Job Tenure		Exp.=Tenure	
	OLS	IV	OLS	IV	P-values	
	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	0.051***	0.037**	0.054***	0.053***	0.762	0.059
	(0.012)	(0.013)	(0.010)	(0.011)		
AFQT×Exper/120	0.036***	0.052***			0.071	0.208
	(0.011)	(0.011)				
AFQT×Tenure/120			0.065***	0.067***		
			(0.018)	(0.015)		
Black	-0.046*	-0.055*	-0.127***	-0.098***	0.000	0.007
	(0.021)	(0.023)	(0.019)	(0.019)		
Black×Exper/120	-0.042*	-0.042			0.005	0.979
	(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		Test: Exper=Tenure	
	OLS	IV	OLS	IV	P-values	
	(1)	(2)	(3)	(4)	(5)	(6)
Standard. AFQT	0.123***	0.119***	0.156***	0.133***	0.122	0.472
	(0.026)	(0.027)	(0.024)	(0.022)		
AFQT×Exper./120	0.038	0.038			0.052	0.680
	(0.024)	(0.025)				
AFQT×Tenure/120			-0.036	0.025		
			(0.041)	(0.032)		
Black	0.138**	0.163**	0.104*	0.104*	0.427	0.127
	(0.047)	(0.051)	(0.046)	(0.043)		
Black×Exper/120	-0.091*	-0.129**			0.416	0.585
	(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: see note to Table 3.

compared to the corresponding OLS estimate, but the IV estimate of the coefficient on AFQT-experience interaction is larger. Both of them remain positive and significant. The estimated speed of employer learning is larger when actual experience is treated as endogenous. Similarly, the IV estimate of the *Black* coefficient is negative and statistically significant, whereas the IV point estimate of the *Black*-experience coefficient is very close to its OLS counterpart. These results from specifications (1) and (2) are consistent with the hypothesis that employers learn about worker productivity over time. In contrast to Altonji and Pierret (2001), we find some evidence on race-based statistical discrimination.[14]

If employer learning is asymmetric, learning takes place mostly on job tenure as outside firms have little information regarding a worker's productivity. Specification (3) in Panel 1 estimates equation (6) for the non-college sample using tenure as time measure. The coefficient of 0.054(0.010) on AFQT is significantly positive and close to the coefficient on AFQT when experience is used as time measure in specification (1). The coefficient of 0.065(0.018) on AFQT-tenure interaction term is also positive and significant, consistent with the prediction of the employer learning model. The point estimate of the coefficient on AFQT-tenure interaction in (3) is greater than the estimated coefficient on AFQT-experience interaction in (1), suggesting that the speed of employer learning is faster over job tenure than over work experience. The estimated coefficient on the race dummy indicates that black workers earn wages 12.7 percent lower than white workers when they start a job, and this wage gap decreases insignificantly over workers' tenure. Therefore when tenure is used as time measure, we also find evidence for statistical discrimination on the basis of race for non-college graduates.

In specification (3), we treat job tenure as exogenous. But tenure should be a function of quit and layoff decisions, and it will 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. In specification (4) we use a simple IV 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, we use the deviations of the job tenure variables around their means for the sample observations on a given job match as an instrumental variable. This variable is

[14] Altonji and Pierret (2001) 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 Altonji and Pierret's finding, but his empirical results imply that the pattern might differ across occupations.

by construction uncorrelated with both the individual and job specific unobserved components. Overall the IV estimates are very similar to the OLS estimates.

In columns (5) and (6) of Table 4, we test if the coefficients presented are significantly different when actual experience versus job tenure are used as time measure. Column (5) presents the P-values of the difference between the two OLS specifications in columns (1) and (3). We do not find significant differences between the coefficients for AFQT in these two specifications. In contrast, the coefficient on AFQT interacted with experience 0.036(0.011) is much smaller than that on AFQT interacted with tenure 0.065(0.018), and these two coefficients are significantly different from each other at 10% level. The quantitatively similar initial AFQT coefficients and the larger AFQT-tenure coefficient indicate that employer 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. The coefficients on *Black* using either experience or tenure are both negative and significant, supporting the hypothesis that employers have limited information about the productivity of new workers and statistically discriminate on the basis of race. The P-value rejects the equality of these two *Black* coefficients. One possible explanation is that employers statistically discriminate against non-college black workers more on jobs that require prior experience than on entry-level jobs.

In Panel 2 of Table 4 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 nearly perfectly 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 wage than their white counterparts when they start a job, and this black wage premium declines over time.[15] Arcidiacono et al. (2010) argue that information

[15]The existence of a substantial black wage premium for college graduates is a robust feature of

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.[16] 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.

5 Additional results

5.1 Non-Purely Asymmetric Learning

Our learning-based racial statistical discrimination model assumes that learning is either symmetric or purely asymmetric. Symmetric learning implies that current and outside firms have the same information on workers' productivity. Purely asymmetric learning, in contrast, suggests that only current firms accumulate information about workers' productivity and outside firms receive no information. It is likely that outside firms receive some new information about workers' productivity over time but they have information disadvantage compared to current firms. When employer learning is not purely asymmetric, outside firms could learn some aspects of workers' productivity, so employer learning occurs both over experience and over tenure.

To examine the possibility that employer learning is not purely asymmetric, We analyze how AFQT vary with experience and tenure when both are included in the regression model to distinguish between alternative learning hypotheses.[17] If learn-

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).

[16]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.

[17]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)	(1) = (2) P-values
Standard. AFQT	0.054***	0.046***	0.333
	(0.010)	(0.012)	
AFQT×Tenure/120	0.065***	0.050*	0.174
	(0.018)	(0.022)	
AFQT×Exper/120		0.018	
		(0.013)	
Black	-0.127***	-0.059**	0.000
	(0.019)	(0.021)	
Black×Tenure/120	0.049	0.097*	0.016
	(0.037)	(0.043)	
Black×Exper/120		-0.065**	
		(0.023)	
R^2	0.253	0.277	
No. of obs.	247,140	247,140	

ing 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.

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

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 along with the estimation results we obtained under the assumption of purely asymmetric learning for comparison. The estimation results under the assumption of purely asymmetric learning appear in column (1) of Table 5, and column (2) shows the empirical results applying regression equation (7) where both tenure and experience interactions are considered. The fact that the coefficient on AFQT-experience interaction term is not statistically different from zero provides empirical evidence in favor of purely asymmetric learning. Unlike AFQT-experience interaction, the coefficient on AFQT-tenure interaction in column (2) remains large in magnitude and statistically significant. The P-values suggest that there are no statistical differences between the estimated coefficients on AFQT and AFQT-tenure interaction in specifications (1) and (2). These results indicate 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. Furthermore, statistical discrimination on the basis of race increases over time for outside firms. 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 Match Quality and Job Mobility

Our empirical results reveal that whenever non-college black workers start a new job, employers pay them significantly less than their white counterparts even conditional on their AFQT scores. This finding is consistent with the view that employers use race as a cheap source of information to infer workers' productivity and thus statistically discriminate against black workers. But one possible explanation for the racial wage gap is that the match quality between non-college black workers and their employers is worse than the match quality between non-college white workers and their employers.

If the observed racial wage gap reflects racial differences in match quality rather than statistical discrimination, we would expect that non-college black workers generally switch firms more frequently than do non-college white workers. Poorly matched black workers are expected to be more likely to move between firms both voluntarily and involuntarily. We test this alternative explanation by estimating the following probit model that examines the effect of race on workers' probability of job change:

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

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. The main coefficient of interest is β_1. If the racial wage gap in the non-college market is attributed to match quality differences across racial groups, we would expect β_1 to be positive for the non-college sample. That is, black non-college workers change job more frequently than their white counterparts because of the inferior match quality. On the other hand, if the racial wage gap comes from statistical discrimination, job mobility rate should be lower for black workers. Black workers have less incentive to switch jobs to avoid being discriminated by their new employers especially when employer learning is asymmetric. For comparison, 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 6. 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 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 0.63% less likely to make a month-to-month job change compared to their white counterparts. This result does not provide support to the alternative interpretation that the racial wage gap for non-college graduates reflects racial differences in firm-worker match quality. Instead, it provides further evidence for the asymmetric learning and statistical discrimination model established in this paper. 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. Specification (3) further controls for urban residence, region of residence and year effect showing no significant changes.

The results for college graduates appear in columns (4) to (6) in Table 6. Black college workers seem to change jobs more frequently than white workers based on the estimates in column (4), 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 (5), 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.

In short, the results of the probit model provide several pieces of evidence in favor of our main findings. The fact that black non-college graduates change job less frequently than do white non-college graduates strengthens our previous finding that black workers are statistically discriminated in the non-college market. Furthermore, there is no statistically significant racial differences in job mobility patterns in the college market, confirming that employers almost perfectly observe the productivity of college-educated workers and therefore have weak incentive to discriminate against blacks.

5.3 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, there may be explanations other than learning-based racial statistical discrimination for the observed racial wage differences. One possible 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 entire menu of workers' career paths. What appears to be evidence of racial statistical discrimination could be attributed to differences in job sorting by black and white workers.[18]

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

Table 6: Probit Estimates on Racial Difference in Job Change Probabilities

	Non-College Graduates			College Graduates		
	(1)	(2)	(3)	(4)	(5)	(6)
Black	-0.000	-0.078***	-0.082***	0.090**	0.026	0.025
	(0.011)	(0.012)	(0.012)	(0.028)	(0.035)	(0.036)
Education	0.002	0.022***	0.010**	-0.003	0.005	0.001
	(0.003)	(0.004)	(0.004)	(0.008)	(0.008)	(0.009)
Experience	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Standard. AFQT		-0.086***	-0.073***		-0.071***	-0.075***
		(0.007)	(0.007)		(0.019)	(0.019)
Urban and region dummies	No	No	Yes	No	No	Yes
Year dummies	No	No	Yes	No	No	Yes
Log pseudolikelihood	-133,374	-133,034	-131,993	-24,702	-24,671	-24,527
No. of obs.	848,484	848,484	848,484	181,452	181,452	181,452

Notes: The numbers in parentheses are White/Huber standard errors that account for multiple observations per person. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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 (5) and (6) as an additional control, and repeat the empirical analysis separately for non-college graduates and college graduates.[19] The regression results are presented in columns (1) and (4) of Table 7. 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 insignificantly with tenure, 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.[20]

We also explore the role of sector allocation by examining the effect of initial

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

[20]The results shown in Table 7 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.

industry on the observed racial wage gap.[21] With initial industry included as additional controls in columns (2) and (5) in Table 7, we repeat the empirical analysis for the two educational groups of interest. The results closely resemble those without the inclusion of initial industry in Table 4. Finally we control for initial occupation and industry simultaneously in columns (3) and (6) in Table 7, 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.[22]

6 Conclusion

In this paper, we combine elements of both employer learning and statistical discrimination theories to develop a learning-based racial statistical discrimination model. We formulate 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 and base their payment more on true productivity as they acquire more information. For college graduates, employers are able to learn most of their productivity upon initial entry into the labor market, and very little learning occurs afterwards. 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. By comparing the coefficients of interests from wage regressions using either general experience or job tenure as time measure, we find empirical evidence for asymmetric learning in the non-college market. As non-college workers spend more time in the labor market, their current firms accumulate more information and gradually learn about their productivity, but outside firms are insulated from the learning process and unable to update their expectations about workers' productivity.

Following the literature, this paper assumes that the AFQT score is a correlate of productivity that is unobserved by employers. If the racial AFQT differences are

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

[22] Table 77presents 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.

Table 7: 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

Notes: The experience measure is actual work experience in months. All specifications control for year effects, urban residence, and region of residence. Specifications with experience also control for a quadratic term in actual 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.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the results of pre-market discrimination or racial bias in testing, then our results should be interpreted as a lower bound estimate of racial discrimination in the labor market.

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[23]. Pinkston (2009) applies the framework of employer learning to test this hypothesis, and his estimation results provide evidence supporting this view.[24] Our learning-based racial statistical discrimination model assumes that the signals sent by workers from different racial groups are equally informative. Statistical discrimination arises because employers have the perception that the average productivity of black workers is lower than that of white workers. 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 racial statistical discrimination.

Finally, we note that our paper, as all of the related literature, is not designed to measure to what extent the notable and persistent racial wage differences[25] are due to statistical versus "taste-based" discrimination (in the sense of Becker (1971)). Disentangling the nature of group inequality remains an important avenue of future research.

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

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

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

References

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

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 2001, *116* (1), 313–350. 2, 13, 14

Altonji, Joseph G and Rebecca M Blank, "Race and gender in the labor market," *Handbook of labor economics*, 1999, *3*, 3143–3259. 30

Altonji, Joseph G. and Robert A. Shakotko, "Do Wages Rise with Job Seniority?," *The Review of Economic Studies*, 1987, *54* (3), 437–459. 20

Arcidiacono, Peter, "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524. 22

_ , 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, 13, 17, 21

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, 30

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. 30

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. 30

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. 26

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. 30

_ **and Jee-Yeon K Lehmann**, "Racial discrimination in the labor market: Theory and empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006. 2, 30

_ **and Michael Manove**, "Education and labor market discrimination," *American Economic Review*, 2011, *101* (4), 1467–96. 22

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. 30

Mansour, Hani, "Does employer learning vary by occupation?," *Journal of Labor Economics*, 2012, *30* (2), 415–444. 2, 20

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. 22, 30

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, 30

Schönberg, Uta, "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 2007, *25* (4), 651–691. 4, 22

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