Testing for Statistical Discrimination with Asymmetric Employer Learning

Suqin Ge  
*Virginia Tech*

Beibei Zhu  
*Virginia Tech*

July, 2012

Abstract

This paper combines elements of both employer learning and screening discrimination theories to develop a learning-based racial statistical discrimination model. We propose a framework that nests both symmetric and asymmetric employer learning, and we derive testable hypotheses on racial statistical discrimination with asymmetric employer learning. Testing the model with data from the NLSY79, we find that employers statistically discriminate against black workers in the high school market where learning appears to be mostly asymmetric. For college graduates, employers directly observe most of the productivity of potential employees and learn very little over time. A series of sensitivity tests provide further support for our main findings.

Keywords: employer learning, statistical discrimination, racial discrimination

*JEL code: J71, D82, J31.*

*Contact information: Ge, Department of Economics, Virginia Tech, Email: ges@vt.edu. Zhu, Department of Economics, Virginia Tech, Email: ivyzhu@vt.edu.*
1 Introduction

In a world where the productivity of labor force participants is difficult to directly observe, employers have to rely on some observable characteristics of workers to infer their unobservable productivity. When hiring, employers could only use information contained in resumes, letters of recommendation and job interviews to evaluate the productivity of potential workers. As workers spend more time in the labor market, new information about their performances become available, and employers could learn about their productivity from newly acquired information. The employer learning literature has confirmed that employers make use of some easily observable characteristics, such as education, to distinguish among inexperienced workers at the time of hire, and their pay will be more dependable on productivity as they accumulate more information.

Some questions arise naturally from this employer learning process. Do employers statistically discriminate against blacks if black workers belong to a group with lower average productivity? Is the process of employer learning that takes place after hiring symmetric or asymmetric? This article attempts to provide an answer to these two questions.

There is an extensive literature on the issue of racial statistical discrimination. Empirical research on racial wage gap documents a notable black-white wage gap for male workers. Lang and Manove (2011) find that a substantial racial wage gap emerges when controlling for AFQT score and education, and this wage gap could not be explained by the differences in the quality of schools attended by blacks and by whites. Their findings support the view that statistical discrimination is one source of
the black-white wage gaps. On the theoretical side, statistical discrimination models are developed to explore the implications of imperfect information about worker’s productivity. ¹ The crucial assumption in statistical discrimination literature is that imperfectly informed employers rationally use group statistics to infer information about productivity.

This article aims to test for statistical discrimination on the basis of race from the perspective of employer learning. We combines elements of both employer learning and screening discrimination theories to formulate a learning-based racial statistical discrimination model. Screening discrimination models that originated from Phelps (1972) are one main branch of the statistical discrimination literature, and attribute discriminatory outcomes to differential observability of productivity. ² Our model borrows insights from screening discrimination literature, and assumes that the average productivity of black workers is lower than that of white workers. ³ Racial statistical discrimination arises because employers know the racial differences in pro-

¹Unlike taste-based models where employers have prejudice against black workers, statistical discrimination models assume that employers rationally uses race to infer worker’s productivity in a world of imperfect information.

²The other main branch is the rational stereotyping models that originated from Arrow (1973), which assume that employer’s negative beliefs about the quality of black workers are self-fulfilling.

³This version of statistical discrimination is discussed in Phelps (1972) and Aigner and Cain (1977). Two groups differ with respect to the average productivity but not with respect to the variance of error term. Sattinger (1998) is an example that builds on an extension to this version of statistical discrimination. Workers are assumed to be homogenous in productivity, but their quitting behaviors differ across groups. One group has a greater proportion of workers whose quit-rate is high. Firms observe quit rates imperfectly and thus rationally set unequal employment criteria or unequal interview rates.
ductivity. Race is a correlate of productivity, and thus is incorporated into the model of employer learning. Race could work as either an easy-to-observe variable or a hard-to-observe variable depending on whether or not employers initially rely on race to predict productivity.

The benchmark employer learning model was developed by two landmark articles, Farber and Gibbons (1996) and Altonji and Pierret (2001). Altonji and Pierret (2001) establish the model of employer learning and statistical discrimination, and much of the subsequent literature on employer learning is guided by the framework set out in this influential study. Many recent studies provide empirical evidence for this model, and further find that the employer learning process differs across educational groups.

Some studies on employer learning explore the issue of learning-based racial statistical discrimination, but all these studies examine racial statistical discrimination under the assumption that employer learning is symmetric, that is, both current and outside firms share the same information about worker’s productivity. Earlier literature on employer learning usually build around the assumption of symmetric learning, however, the empirical studies on whether learning is symmetric or asymmetric so far fail to reach a consensus.

Another main contribution of this article is to distinguish between symmetric and asymmetric learning in the labor market. The issue of racial statistical discrimination is addressed under the assumption of different types of employer learning. To the best of my knowledge, there are no studies that investigate learning-based racial statistical

---

4Farber and Gibbons (1996) develop their empirical predictions for wage levels while Altonji and Pierret (2001) work with log wages.  
discrimination under asymmetric learning. The nature of employer learning has
great influence on racial statistical discrimination. If learning is asymmetric, there
might be more scope for racial statistical discrimination. The article attempts to fill
the gap in the current employer learning literature by developing a testable model
that nests both learning hypotheses and testing for racial statistical discrimination.
We build on some previous work, particularly Altonji and Pierret (2001) and Lange
(2007), set up a framework that nests both learning hypotheses, and propose a
new way to distinguish between two types of employer learning. We differentiate
between general experience and job tenure to explore whether employer learning is
symmetric or asymmetric. Symmetric learning implies a continuous learning process
over a worker’s general experience. If employer learning is asymmetric, however,
the learning process should mainly take place over job tenure instead of general
experience.

In this article, we use the 2008 release of National Longitudinal Survey of Youth
1979 (NLSY79) to test racial statistical discrimination and to distinguish between two
types of employer learning. After confirming Altonji and Pierret (2001)’s finding of
employer learning and statistical discrimination on the basis of education, we conduct
the empirical analysis for both the full sample and the two educational groups, high
school graduates and college graduates. 6 For the full sample, we could not draw
any definite conclusion because the empirical analysis gives conflicting results. The
results for high school graduates provide evidence in favor of statistical discrimination
by race and asymmetric learning. The college sample reveals a distinct pattern

---

6The employer learning literature (e.g., Arcidiacono, Bayer, and Hizmo (2010)) find that different educational groups are likely to be associated with distinct patterns of employer learning.
of employer learning. In the college market, the key aspects of productivity are directly observed upon initial entry and thus little learning takes place subsequently, supporting the main findings reported in Arcidiacono, Bayer, and Hizmo (2010). A series of sensitivity tests demonstrate that our main results are robust to alternative explanation of the data. The process of employer learning for high school graduates tends to be more like purely asymmetric, and the black-white wage gap can not be attributed to racial differences in match quality or the skill level of jobs.

The structure of the article 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 empirical specifications used in the empirical analysis. The estimation results are reported in Section 4, and Section 5 show their robustness to alternative interpretations. Section 6 concludes and discusses future work.

2 The Learning-based Racial Statistical Discrimination Model

In this section, we develop a learning-based racial statistical discrimination model, and derive testable predictions under different assumptions. Much of the model is based on Altonji and Pierret (2001) and Lange (2007). Altonji and Pierret (2001) first propose the theory of employer learning and statistical discrimination, and their framework is widely applied in the employer learning literature. Lange (2007) builds on Altonji and Pierret (2001) and constructs a framework to formally estimate the speed of employer learning.
The model developed in this section investigates how the wage coefficients on correlates of productivity evolve over time, and examines the time trend of wage coefficients to test whether or not employers statistically discriminate against black workers. Furthermore, our model provides predictions that could distinguish between symmetric and asymmetric learning.

2.1 The Employer Learning Model

Suppose workers first enter the labor market at time $t = 0$ and stay with their first employers during the time period $t$ under consideration.

Following the standard literature, we specify that worker $i$’s log productivity at time $t_i$ depends linearly on a set of variables $(s_i, q_i, z_i, \eta_i)$ and a polynomial in time $\tilde{H}(t_i)$:

$$\chi_{i,t} = \tilde{\chi}(s_i, q_i, z_i, \eta_i) + \tilde{H}(t_i)$$

$$= \gamma s_i + \alpha q_i + \lambda z_i + \eta_i + \tilde{H}(t_i).$$

(1)

The variables $s_i$ represent the information available to both employers and researchers, such as education. The variables $q_i$ describe the information available to employers but not used or observed by researchers, such as information obtained through job interviews. The variables $z_i$ are correlates of productivity that are available to researchers but not to employers. Much of the literature assume that AFQT score is such a correlate of productivity, we also make this assumption. The variables $z_i$ are normalized so that all the elements of $\lambda$ are positive. The variables $\eta_i$ denote the information that is unobserved by both employers and researchers. Finally, $\tilde{H}(t_i)$ denotes the relation between log productivity and time, and is assumed to be inde-
dependent of \((s_i, q_i, z_i, \eta_i)\). Therefore, human capital accumulation is incorporated into the model through experience profiles of productivity, \(\bar{H}(t_i)\), but is assumed to have no influence on the time paths of the impacts of \(s_i\) and \(z_i\) on log productivity. The variation of the time gradient of log wages with \(s_i\) and \(z_i\) is interpreted as the outcome of an employer learning process about worker’s productivity. The subscript \(i\) will be suppressed from now on.

2.1.1 No Statistical Discrimination by Race

If employers obey the law and do not use race as information, then the information used by employers at time \(t = 0\) is \((s, q)\). The black indicator \(\text{Black}\), which takes \(1\) if the worker is black and \(0\) otherwise, can be thought of as a component of the correlates of productivity \(z\), which is unobservable to employers.

Employers observe \((s, q)\) but not \((z, \eta)\), and form the conditional expectation of \((z, \eta)\) on the information available. We assume that \((s, q, z, \eta)\) are jointly normally distributed. Although employers can not directly observe \(z\), it is assumed that \(\bar{z} = E(z|s, q)\) is common knowledge. Employers use the average information to predict \(z\) by the following linear relation:

\[
\begin{align*}
    z &= E(z|s, q) + \nu = \bar{z} + \nu. \\
\end{align*}
\]  

(2)

Because of the normality assumption the expectation of \(\eta\) conditional on \((s, q)\) is also linear in \((s, q)\):\(^7\)

\[
\begin{align*}
    \eta &= E(\eta|s, q) + e = \alpha_1 s + e. \\
\end{align*}
\]  

(3)

\(^7\)We define \(\eta\) and coefficient \(\alpha\) on \(q\) in (1) so that the mean of \(\eta\) does not depend on \(q\).
Substituting (2) and (3) in (1), we can express the initial log productivity as a linear function of the information available to employers at time $t = 0$:

$$\chi = (\gamma + \alpha_1)s + \alpha q + \lambda \overline{z} + \lambda \nu + e + \tilde{H}(0)$$

$$= E[\tilde{\chi}|s, q, \overline{z}] + (\lambda \nu + e) + \tilde{H}(0). \quad (4)$$

As workers accumulate more labor market experience, new information about worker’s job performance become available to employers. In each period $t$, employers obtain a noisy signal $y_t$ of $\tilde{\chi}$:

$$y_t = \tilde{\chi} + \epsilon_t, \quad (5)$$

where the noise $\epsilon_t$ is assumed to be uncorrelated with all other variables and is independently, identically and normally distributed with a variance of $\sigma^2_{\epsilon}$. At time $t$, employers could observe a $t$ dimensional vector of measurements, $Y^t = (y_1, \ldots, y_t)$.

The normality assumption simplifies the process of updating employer expectations about $\tilde{\chi}$. When workers first start their jobs, that is, at time $t = 0$, the prior mean of employers’ beliefs about $\tilde{\chi}$ is

$$\mu_0 = E[\tilde{\chi}|s, q, \overline{z}] = (\gamma + \alpha_1)s + \alpha q + \lambda \overline{z} \quad (6)$$

At time $t > 0$, employers receive a signal of worker’s log productivity $y_t$ and then update their beliefs. Assume that the variance of $\tilde{\chi}$ conditional on initial information $(s, q, \overline{z})$ is $\sigma^2_0$. It is the variance of the initial expectation error $(\lambda \nu + e)$. Because of the normality assumption, the posterior distribution of employers’ beliefs about $\tilde{\chi}$ at
time $t$ is normal with mean $\mu_t$ and variance $\sigma_t^2$,

$$\mu_t = E[\tilde{\chi}|s,q,\bar{z},Y^t] = \frac{\sigma^2}{t\sigma_0^2 + \sigma_e^2}\mu_0 + \frac{t\sigma_0^2}{t\sigma_0^2 + \sigma_e^2}\left(\frac{1}{t} \sum_{t=1}^{t} y_t\right)$$

$$= (1 - \theta_t) \mu_0 + \theta_t \left(\frac{1}{t} \sum_{t=1}^{t} y_t\right)$$  \hspace{1cm} (7)

where

$$\theta_t = \frac{t\sigma_0^2}{t\sigma_0^2 + \sigma_e^2}$$  \hspace{1cm} (8)

and

$$\sigma_t^2 = \frac{\sigma_0^2\sigma_e^2}{t\sigma_0^2 + \sigma_e^2}. $$  \hspace{1cm} (9)

The learning parameter $\theta_t$ lies in the interval $[0,1]$, and converges from 0 to 1 as $t$ increases.

Therefore, the expected productivity of a worker at time $t$ is

$$E[\chi|s,q,\bar{z},Y^t] = (1 - \theta_t) \mu_0 + \theta_t \left(\frac{1}{t} \sum_{t=1}^{t} y_t\right) + \tilde{H}(t)$$ \hspace{1cm} (10)

Employers set the wage equal to the expected productivity conditional on the information available $(s,q,\bar{z},Y^t)$:

$$W(s,q,\bar{z},Y^t) = E[exp(\chi)|s,q,\bar{z},Y^t]$$ \hspace{1cm} (11)

Because of the normality assumption, the distribution of $\chi$ conditional on $(s,q,\bar{z},Y^t)$ is normal. We use the property of a log normal distribution, $E[exp(\chi)|s,q,\bar{z},Y^t] = exp\left(E[\chi|s,q,\bar{z},Y^t] + \frac{1}{2}\sigma_t^2\right) = exp\left(E[\chi|s,q,\bar{z},Y^t] + \tilde{H}(t) + \frac{1}{2}\sigma_t^2\right)$. The expectation error at time $t$ does not depend on $(s,q,\bar{z},Y^t)$, so $\frac{1}{2}\sigma_t^2$ is constant with respect to...
\( (s, q, \bar{z}, Y^t) \). Hence, we define \( H(t) = \tilde{H}(t) + \frac{1}{2} \sigma_1^2 \). Using equation (11) and taking logs, we obtain the following expression for log wages:

\[
\begin{align*}
    w(s, q, \bar{z}, Y^t) &= (1 - \theta_t) \mu_0 + \theta_t \left( \frac{1}{t} \sum_{t=1}^{t} y_t \right) + H(t) \\
    &= (1 - \theta_t) \mu_0 + \theta_t \left( \frac{1}{t} \sum_{t=1}^{t} y_t \right) + H(t) \\
    &= (1 - \theta_t) \mu_0 + \theta_t \sum_{t=1}^{t} y_t + H(t) \quad (12)
\end{align*}
\]

Equation (12) gives the log wages conditional on the information available to employers \((s, q, \bar{z}, Y^t)\). However, researchers observe \((s, z, \bar{z}, t)\) instead of \((s, q, \bar{z}, Y^t)\). We need to express the log wages as a function of \((s, z, \bar{z}, t)\) rather than \((s, q, \bar{z}, Y^t)\). We define the linear projections of \((q, \eta)\):

\[
\begin{align*}
    q &= \gamma_1 s + u_1 \\
    \eta &= \gamma_2 s + u_2
\end{align*}
\]

Therefore, the linear projection of log wages conditional on the information available to researchers \((s, z, \bar{z}, t)\) is given by

\[
\begin{align*}
    w_t &= E^* \left[ w(s, q, \bar{z}, Y^t) \mid s, z, \bar{z}, t \right] \\
    &= (1 - \theta_t) E^* [\mu_0 \mid s, z, \bar{z}, t] + \theta_t E^* [\tilde{\chi} \mid s, z, \bar{z}, t] + H(t). \\
    &= (1 - \theta_t) E^* [\mu_0 \mid s, z, \bar{z}, t] + \theta_t E^* [\tilde{\chi} \mid s, z, \bar{z}, t] + H(t). \\
    &= (1 - \theta_t) E^* [\mu_0 \mid s, z, \bar{z}, t] + \theta_t E^* [\tilde{\chi} \mid s, z, \bar{z}, t] + H(t). \quad (15)
\end{align*}
\]

Equations (6) and (13) imply

\[
E^*[\mu_0 \mid s, z, \bar{z}, t] = \lambda z + (\gamma + \alpha_1 \gamma) s, \quad (16)
\]

and equations (1), (13) and (14) imply

\[
E^*[\tilde{\chi} \mid s, z, \bar{z}, t] = \lambda z + (\gamma + \alpha_1 \gamma + \gamma_2) s. \quad (17)
\]

Substituting (16) and (17) into (15) and rearranging terms results in

\[
w_t = \lambda \{(1 - \theta_t) \bar{z} + \theta_t z\} + \chi_t s + H(t) \quad (18)
\]
where
\[ \chi_t = (1 - \theta_t)(\gamma + \alpha_1 + \alpha_1) + \theta_t(\gamma + \alpha_1 + \gamma_2). \] (19)

The weights that employers place on \( \bar{z} \) and \( z \) are given by \((1 - \theta_t)\) and \( \theta_t \), respectively. \( \theta_t \) is the learning parameter that reflects the process of employer learning.

If there is no statistical discrimination on the basis of race, employers do not rely on race to predict the productivity of new workers. However, significant productivity differences exist among different racial groups. Consequently, we assume that the correlates of productivity \( z \) consists of two components, AFQT score and black indicator \( Black \). The correlates of productivity \( z \) can be expressed as
\[ z = \tau_1 AFQT + \tau_2 Black. \] (20)

Most literature assumes that AFQT score is a positive correlate of productivity, and we assume that as well. In the NLSY79 sample, the mean of the standardized AFQT score for blacks is about one standard deviation below the mean for whites. The fact that the average ability of black workers is lower than that of whites implies that the black indicator \( Black \) is negatively correlated with the productivity. Hence, \( \tau_1 > 0 \) and \( \tau_2 < 0 \).

Plugging (20) into (18), the wage equation can be rewritten as
\[ w_t = \theta_t \{ \lambda_1 AFQT + \lambda_2 Black \} + \chi_t s + \text{other} \] (21)

where
\[ \text{other} = H(t) + \lambda(1 - \theta_t)\bar{z}. \] (22)
At the time of initial hire, the learning parameter $\theta_t$ equals 0, and converges to 1 over time. Therefore, equation (21) implies that the coefficient on the positive correlate of productivity $AFQT$ increases from 0 to $\lambda \tau_1$ and that on $Black$ decreases from 0 to $\lambda \tau_2$ over time since $\lambda \tau_1 > 0$ and $\lambda \tau_2 < 0$. The time paths of the coefficients on $AFQT$ and $Black$ are shown in Figure 1. This is the basis for Proposition 1.

**Proposition 1** In the case of no racial statistical discrimination, the coefficient on $Black$ is initially 0 and decreasing in $t$.

![Diagram of coefficient paths](image)

Figure 1: No Racial Statistical Discrimination

### 2.1.2 Statistical Discrimination by Race

If employers know the average productivity for each race group and use race as a cheap source of information, then they will statistically discriminate against black workers at the time of initial hire. In this case, the black indicator $Black$ acts as an easily observable variable to employers. The information available to employers at time $t = 0$ now is $(s, q, Black)$, and employers use the black indicator $Black$ to
predict correlates of productivity $z$ by the following linear relation:\(^8\)

$$z = E(z|s, q, \text{Black}) + \nu = \bar{z} + \delta \text{Black} + \nu$$  \hspace{1cm} (23)

where $\delta < 0$ resulting from the fact that black workers are associated with lower average productivity. Employers know the average differences between blacks and whites, and hence $E(z|s, q, \text{Black})$ is written as a function of $\bar{z}$ and $\text{Black}$.

Employers use the racial information to predict the productivity of workers when they first enter the labor market, so $\text{Black}$ is a component of the easy-to-observe characteristics now. The hard-to-observe correlates of productivity $z$ only include AFQT score and can be expressed as

$$z = \tau_3 \text{AFQT}$$  \hspace{1cm} (24)

where $\tau_3 > 0$ because AFQT score is positively correlated with productivity.

Using (23) and (24), we obtain the following wage equation:

$$w_t = \{\theta_t \lambda \tau_3 \text{AFQT} + (1 - \theta_t) \lambda \delta \text{Black}\} + \chi_t s + \text{other}.$$  \hspace{1cm} (25)

The learning parameter $\theta_t$ is bounded between 0 and 1, and increases to 1 as the worker’s career progresses. Hence, the equation (25) implies that the coefficient on $\text{AFQT}$ increases from 0 to a positive value, $\lambda \tau_3$, and that on $\text{Black}$ is initially negative, $\lambda \delta$, and rises to 0 during the employer learning process. The time paths of the coefficients on $\text{AFQT}$ and $\text{Black}$ are shown in Figure 2.

---

\(^8\)One implicit assumption is that the signals of different racial groups are equally informative, that is, $\sigma_{sB} = \sigma_{sW} = \sigma_s$. Most screening discrimination models build around the crucial assumption that the signals employers receive from black workers are noisier than that from white workers. In our model, statistical discrimination arises because of the assumption that the average productivity of black workers is lower than that of white workers.
Proposition 2 In the case of racial statistical discrimination, the coefficient on Black is initially negative and increasing in t.

![Graph showing the coefficient on AFQT and Black over time](image)

Panel A. The coefficient on AFQT

\[ \beta_{AFQT} = \lambda t \eta_t \]

Panel B. The coefficient on Black

\[ \beta_{Black} = \lambda \delta (1 - \theta_t) \]

Figure 2: Racial Statistical Discrimination

2.2 Symmetric Learning

The benchmark employer learning model developed by Farber and Gibbons (1996) and Altonji and Pierret (2001) assume that employer learning is symmetric, that is, all firms share the same information about worker’s productivity. If employer learning is symmetric, both current firms and outside firms learn about worker’s productivity over time, and thus the employer learning process occurs over the worker’s general labor market experience. We will derive the implications of symmetric learning first, and then move on to the case of asymmetric learning.

We use \( X \) denotes worker’s general experience and \( T \) denotes worker’s job tenure with a specific firm. It is worth noting that the theory of human capital also distinguishes between two types of experience, general experience and firm-specific ex-
The employer learning theory differs from the human capital theory in the sense that it focuses on the time paths of the effects of easy-to-observe and hard-to-observe variables on productivity.

Consider the case where the worker changes $n$ jobs in his lifetime. The lengths of job tenure for each job are $T_1$, $T_2$, \ldots, $T_n$, and $T_1 + T_2 + \cdots + T_n = a$, where $a$ is the length of general experience for this worker. At time $t = T_1$, the worker moves to a new job, and has $T_1$ units of general experience and 0 unit of tenure with his new employer. At time $t = T_1 + T_2$, the worker has spent $T_2$ units of time on his new job, with $T_1 + T_2$ units of general experience and $T_2$ units of tenure. The time paths of the learning parameter $\theta_t$ over general experience $X$ and over tenure $T$ are shown in Figure 3 where we assume $n = 3$ to simplify the matter.

When the worker’s career starts at time $t = 0$, the learning parameter $\theta_t(t = 0)$ equals 0. At that time, the worker has zero working experience, that is, $\theta^X(X = 0) = \theta_t(t = 0) = 0$. Whenever a worker starts a new job, however, it is very likely that he has some previous experience in the labor market, implying that $\theta_T(T = 0) > \theta_t(t = 0) = 0$. For example, the worker has $T_1$ units of general experience when he starts a new job at $t = T_1$, and $\theta_T(T = 0) = \theta^X(X = T_1) > 0$ as shown in Figure 3. Therefore, $0 = \theta^X(X = 0) \leq \theta^T(T = 0)$.

Because $\theta_t' = \frac{\sigma^2 \sigma_t^2}{(\sigma_0^2 + \sigma_t^2)^2} > 0$ and $\theta_t'' = \frac{-2 \sigma_0^4 \sigma_t^2}{(\sigma_0^2 + \sigma_t^2)^3} < 0$, the learning parameter $\theta_t$ increases at a decreasing rate, implying that the speed of employer learning declines over time. As a result, the speed of learning over general experience is faster than

---

\textsuperscript{9}Human capital theory argues that general experience increases the worker’s productivity in both the current firm and outside firms but firm-specific experience only increases the worker’s productivity in the current firm.
that over tenure. \(^{10}\) Let \(K\) denotes the speed of employer learning, then \(K^X > K^T\).

**Proposition 3** In the case of symmetric learning, the initial learning parameter \(\theta_t\) over general experience is not greater than that over tenure, and the speed of employer learning is faster over general experience than that over tenure.

![Figure 3: Symmetric Learning](image)

The time path of the coefficients on \(AFQT\) shown in Panel A of Figure 1 and Figure 2, \(\beta_{AFQT} = \lambda \tau_s \theta_t\) where \(\lambda > 0\) and \(\tau_s > 0\), \(^{11}\) is entirely determined by the learning parameter \(\theta_t\). The learning parameter \(\theta_t\) gradually rises from 0 to 1 as employers updating their expectations about worker’s productivity. We could derive the following predictions about the coefficients on \(AFQT\) from Proposition 3.

1. The initial coefficient on \(AFQT\) using general experience measure is not greater than that using tenure measure, that is, \(0 = \beta_{AFQT}^X(X = 0) \leq \beta_{AFQT}^T(T = 0)\).

\(^{10}\)One exception is that the worker stays with one firm in his entire career life. In that case, the speed of employer learning over general experience equals that over tenure.

\(^{11}\)If employers do not statistically discriminate against black workers, then \(\beta_{AFQT} = \lambda \tau_1 \theta\). However, if statistical discrimination on the basis of race does exist in the labor market, then \(\beta_{AFQT} = \lambda \tau_3 \theta\). As discussed earlier, both \(\tau_1\) and \(\tau_3\) are positive.
2. The coefficient on $AFQT$ increases over time, so $K^X_{AFQT} > K^T_{AFQT} > 0$.

The variation of the coefficient on black indicator $Black$ depends on whether race acts as an easy-to-observe variable such as eduction or as an hard-to-observe variable such as AFQT score. If employers use race as information to predict productivity, race enters into the employer learning model as an hard-to-observe variable, and the learning-based racial statistical discrimination model developed in this article predicts a narrowing racial wage gap. In contrast, if employers do not rely on race, race can be understood as an easy-to-observe variable, and we could expect to observe a widening racial wage gap.

If employers obey the law, and do not statistically discriminate against black workers, then the time path of the coefficient on $Black$ is $\beta_{Black} = \lambda \tau_2 \theta_t$ where $\tau_2 < 0$ reflecting the negative correlation between $Black$ and productivity. Proposition 1 and Proposition 3 leads to the following implications about the coefficients on $Black$.

1. The initial coefficient on $Black$ is zero using general experience measure, but negative using tenure measure. That is, $0 = \beta^X_{Black}(X = 0) \geq \beta^T_{Black}(T = 0)$.

2. The coefficient on $Black$ declines over time, so $K^X_{Black} < K^T_{Black} < 0$.

If statistical discrimination on the basis of race does exist in the labor market, the time path of the coefficient on $Black$ is $\beta_{Black} = \lambda \delta (1 - \theta_t)$ where $\lambda > 0$ and $\delta < 0$. Proposition 3 implies that $1 - \theta^X(X = 0) \geq 1 - \theta^T(T = 0) > 0$. Moreover, the learning parameter $\theta_t$ is increasing in $t$, so the weight on $Black$, $(1 - \theta_t)$, is declining over time. The predictions derived from Proposition 2 and Proposition 3 are as follows.
1. The initial coefficient on Black using general experience is more negative than that using tenure, that is, $\beta^X_{\text{Black}}(X = 0) \leq \beta^T_{\text{Black}}(T = 0) < 0$.

2. The coefficient on Black increases over time, so $K^X_{\text{Black}} > K^T_{\text{Black}} > 0$.

### 2.3 Asymmetric Learning

Many recent studies (Bauer and Haisken-DeNew (2001), Schonberg (2007), Pinkston (2009)) find empirical evidence in favor of asymmetric employer learning, that is, the current firms accumulate more information about the worker’s productivity than do outside firms. If employer learning is asymmetric, the learning process mainly takes place over tenure. As the worker spends more time in the workplace, the current firms observe worker’s performance, and update their expectations about worker’s productivity. The worker’s job performance information in a specific firm is only available to that firm, so outside firms are insulated from the employer learning process.

Consider a worker with $a$ units of general experience. If the worker changes $n$ firms, then the length of his general experience $X$ is the sum of the tenure with each firm, that is, $X = T_1 + T_2 + \cdots + T_n = a$. If the worker does not switch job at all and stays at one firm, then the length of his general experience $X$ is the same as that of tenure, that is, $X = T = a$. The time paths of the learning parameter $\theta_t$ over general experience $X$ and over tenure $T$ are shown in Figure 4 where we simply assume that the worker switches jobs 3 times during the time period under consideration.

---

12Here we discuss the case where employer learning is pure asymmetric. In the section of robustness check, we examine the possibility that employer learning is imperfect asymmetric rather than pure asymmetric.
Because the employer learning process only happens over the worker’s tenure, the learning parameter $\theta_t$ equals zero whenever the worker moves to a new firm, that is, $\theta^T(T = 0) = 0$. When the worker initially enters the labor market and begins his first job at time $t = 0$, both his general experience and tenure equal zero, so $\theta^X(X = 0) = \theta^T(T = 0) = 0$.

In the case of asymmetric learning, outside firms have no access to the rich information about worker’s job performance in other firms, so the employer learning process starts all over again whenever the worker makes a job change. Given a specific time period, therefore, the speed of employer learning is faster over tenure than over general experience, that is, $K^X < K^T$. As shown in Figure 4, $\theta^X(X = a) = \theta^T(T = T_3) < \theta^T(T = a)$. Moreover, $K^X = 0$ whenever the worker moves to a new firm.

**Proposition 4**  
*In the case of asymmetric learning, the initial learning parameter $\theta_t$ over general experience is the same as that over tenure, and the speed of employer learning is slower over general experience than that over tenure.*

We could derive the following predictions from Proposition 1, 2 and 4.

1. The initial coefficients on $AFQT$ and $Black$ are the same either we use general experience or tenure as the time measure. That is, $\beta^X_{AFQT}(X = 0) = \beta^T_{AFQT}(T = 0) = 0$, and $\beta^X_j(X = 0) = \beta^T_j(T = 0)$.

2. The coefficient on $AFQT$ always increases over tenure, and the speed of employer learning is faster using tenure than using experience, that is, $0 \leq K^X_{AFQT} < K^T_{AFQT}$.
Panel A. The time path of $\theta$ over Experience

Panel B. The time path of $\theta$ over Tenure

Figure 4: Asymmetric Learning

3. If employers do not statistically discriminate against black workers, the coefficient on $Black$ is initially zero and decreasing over tenure. That is, $\beta_{Black}^X(X = 0) = \beta_{Black}^T(T = 0) = 0$ and $0 \geq K_{Black}^X > K_{Black}^T$.

4. In the case of statistical discrimination on the basis of race, the coefficient on $Black$ is initially negative and increasing over tenure. That is, $\beta_{Black}^X(X = 0) = \beta_{Black}^T(T = 0) < 0$ and $0 \leq K_{Black}^X < K_{Black}^T$.

3 Data and Empirical Specification

3.1 Data: NLSY79

The empirical analysis is based on the 2008 release of National Longitudinal Survey of Youth (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 for whom the employer learning should matter the most, and its weekly work history data provide the information needed for accurate measurements of both actual experience and tenure.

In selecting the sample, we follow the criteria used in Altonji and Pierret (2001) and Arcidiacono, Bayer, and Hizmo (2010). The empirical analysis is restricted to black and white male workers who have completed at least 9 years of education. We focus exclusively on male workers because the analysis of female workers involves selection issues. We only consider observations after the respondent has made the school-to-work transition. A respondent is considered to have entered the labor market when he has left school for the first time, which is defined as the year of the last enrollment in regular school. Potential experience is the sum of months since the respondent first left school.

For each respondent, we construct the monthly labor market history from the work history data, which contains respondent’s week-by-week labor force status. Each respondent’s weekly work records are transformed into monthly ones. We link all the jobs across different survey years and build a complete employment history for each respondent in the sample. Multiple jobs held at the same time are considered as a new job, and the average wage is used as the wage for the newly constructed job. The average wage is calculated using the hourly rate of pay for each job and the hours per week worked for each job. The deflators from CPI-U released by Bureau of Labor Statistics are used to create the real monthly wage with 1990 as the base year, and all observations with real wages less than $1 or more than $100 are excluded from the analysis. From the newly constructed monthly work history data,
we construct measures of tenure and actual experience. Tenure is computed as the number of months between the start of the job and either the date the job ended or the interview date, and actual experience is calculated as the sum of tenure for each job.  

Following the literature, we use Armed Forces Qualification Test (AFQT), which is administered to the NLSY respondents, as our measure of productivity. The AFQT score provides a summary measure for basic literacy and numeric skills and is thus a correlate of worker’s productivity. To make our measure of productivity comparable to other studies, we standardize the AFQT score to have a zero mean and a standard deviation of one for each three-month age cohort.  

The education variable is the highest grade completed by the respondent at the time of interview. In the empirical analysis, we give special attention to two educational groups, high school graduates and college graduates. High school graduates are defined as workers who have completed 12 years of schooling at the time of interview, and college graduates are workers who have at least 16 of schooling.

Table 1 presents the summary statistics for the NLSY79 sample used in the analysis. It is worth noting that the average AFQT score of black workers is about one standard deviation lower than that of white workers for both the overall sample and the subsamples of two educational group. Consequently, if employers know

\begin{footnote}{In Altonji and Pierret (2001), actual experience is defined as the total number of weeks in which the respondent worked after they leave school for the first time. The actual experience measure we create is more compatible with the tenure measure.}{13}\end{footnote}  

\begin{footnote}{We use the 2006 released AFQT-3 as our measure of AFQT. Within each three-month age group each individual is given a percentile score that ranges between zero and 100. NLS staff recommend using the AFQT-3 because it is renormed controlling for age.}{14}\end{footnote}
the significant productivity differences between black and white workers, they have strong incentives to statistically discriminate on the basis of race. In the next section, we carry out the empirical analysis to examine this issue in detail.

Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>High School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whites</td>
<td>Blacks</td>
<td>Whites</td>
</tr>
<tr>
<td>Real Hourly Wage</td>
<td>1294.36</td>
<td>1019.17</td>
<td>1086.32</td>
</tr>
<tr>
<td></td>
<td>(820.58)</td>
<td>(619.03)</td>
<td>(529.46)</td>
</tr>
<tr>
<td>Potential Experience</td>
<td>132.29</td>
<td>146.35</td>
<td>135.88</td>
</tr>
<tr>
<td></td>
<td>(85.03)</td>
<td>(86.21)</td>
<td>(86.61)</td>
</tr>
<tr>
<td>Actual Experience</td>
<td>111.29</td>
<td>111.84</td>
<td>113.87</td>
</tr>
<tr>
<td></td>
<td>(76.49)</td>
<td>(73.53)</td>
<td>(78.58)</td>
</tr>
<tr>
<td>Job Tenure</td>
<td>47.09</td>
<td>41.06</td>
<td>47.77</td>
</tr>
<tr>
<td></td>
<td>(48.59)</td>
<td>(43.58)</td>
<td>(49.72)</td>
</tr>
<tr>
<td>Education</td>
<td>13.35</td>
<td>12.70</td>
<td>12.00</td>
</tr>
<tr>
<td></td>
<td>(2.39)</td>
<td>(2.01)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Standardized AFQT</td>
<td>0.502</td>
<td>-0.568</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>(0.957)</td>
<td>(0.798)</td>
<td>(0.812)</td>
</tr>
<tr>
<td>No.Individuals</td>
<td>2593</td>
<td>1136</td>
<td>1267</td>
</tr>
<tr>
<td></td>
<td>(59316)</td>
<td>(147)</td>
<td>(688)</td>
</tr>
<tr>
<td>No.Observations</td>
<td>225708</td>
<td>94416</td>
<td>104712</td>
</tr>
</tbody>
</table>

Notes: The data are from the 1979-2008 waves of NLSY79. Standard deviations are in parentheses. Real hourly wages are in cents. Potential experience, actual experience and job tenure are in months.
3.2 Empirical Specification

This section describes how the predictions of the employer learning and statistical discrimination model developed in this article are applied to the data. The learning parameter $\theta_i$ implies that the effects of AFQT score and black indicator $Black$ vary non-linearly with time. To simplify the matter, however, we assume a linear relationship between log wage, AFQT score and black indicator over time in the empirical analysis. To test these predictions, we apply the empirical framework proposed by Farber and Gibbons (1996) and Altonji and Pierret (2001). The wage equation regresses log wages on AFQT score, a black indicator variable which takes 1 if the worker is black and 0 otherwise, education, time interaction terms, and demographic variables:

$$\ln w_{i,t} = \beta_0 + \beta_{AFQT} AFQT_i + \beta_{AFQT,t}(AFQT_i \times t_{i,t}) + \beta_{Black} Black_i + \beta_{Black,t}(Black_i \times t_{i,t}) + \beta_{S} S_i + \beta_{S,t}(S_i \times t_{i,t}) + \beta_{\Omega} \Omega_{i,t} + \epsilon_{i,t}$$  

(26)

where $w_{i,t}$ denotes wage, $Black$ is an indicator variable, $S_i$ denotes education, and $\Omega_i$ is a vector of demographic variables and other controls. All time interaction terms are divided by 120, so the coefficient on interaction terms measures the change in wage during a ten-year period. Our main coefficients of interest are the coefficients on $AFQT$ and $Black$, $\beta_{AFQT}$ and $\beta_{Black}$, and the coefficients on $AFQT$ interacted with time $t$, $\beta_{AFQT,t}$, and $Black$ interacted with time $t$, $\beta_{Black,t}$.

If employer learning is symmetric, the learning process occurs over the general experience regardless of the worker’s job turnover rate. In contrast, asymmetric learning indicates that only current firms learn about worker’s productivity over time, implying a continuous learning process over tenure but not over general experience.
To distinguish between symmetric and asymmetric learning, actual experience $X$ and tenure $T$ are used as time measure $t$ in equation (26), respectively, and the corresponding estimating equations are as follows:

$$
\ln w_{i,t} = \beta_0^X + \beta_{AFQT}X(\text{AFQT}_i \times X_{i,t}) + \beta_{Black}X_{Black_i} \\
+ \beta_{Black,X}(Black_i \times X_{i,t}) + \beta_S^X S_i + \beta_{S,X}(S_i \times X_{i,t}) + \beta_{\Omega}X\Omega_{i,t} + \epsilon_{i,t}
$$

(27)

$$
\ln w_{i,t} = \beta_0^T + \beta_{AFQT}T(\text{AFQT}_i \times T_{i,t}) + \beta_{Black}T_{Black_i} \\
+ \beta_{Black,T}(Black_i \times T_{i,t}) + \beta_S^T S_i + \beta_{S,T}(S_i \times T_{i,t}) + \beta_{\Omega}T\Omega_{i,t} + \epsilon_{i,t}
$$

(28)

If employer learning is symmetric, both current firms and outside firms learn about worker’s productivity. Firms could learn a new worker’s productivity through their past experience, so $\beta_{AFQT}^T$ is larger than $\beta_{AFQT}^X$. The impacts of AFQT score and black indicator on log wages are assumed to vary linearly with experience or tenure, so $\beta_{AFQT,t}$ and $\beta_{Black,t}$ do not vary whether experience or tenure measure is used. The effect of AFQT score on log wage increases over time, so $\beta_{AFQT,t}$ is positive. If there is no statistical discrimination on the basis of race, then $\beta_{Black}^X$ is zero, but $\beta_{Black}^T$ is very likely to be negative. $\beta_{Black,X}^T$ equals $\beta_{Black,T}^T$, and both are negative, implying a widening racial wage gap. On the contrary, racial statistical discrimination predicts that $\beta_{Black}^X$ should be equal to or more negative than $\beta_{Black}^T$, and both $\beta_{Black,X}^X$ and $\beta_{Black,T}^T$ are positive, reflecting a narrowing racial wage gap.

On the other hand, the initial coefficients on AFQT and Black are the same whether we use the experience or tenure measure under the assumption of asymmetric learning. The reason is that a worker without any experience must have a tenure of zero. What is more, the speed of employer learning is faster over tenure than
over experience because the employer learning process occurs only over tenure. As a result, tenure interaction terms are larger in magnitude than experience interaction terms.

Depending on the nature of employer learning and the existence of racial statistical discrimination, there are four possible combinations, and their distinct predictions are summarized in Table 2. The empirical analysis in the next section will rely on these predictions to distinguish between two types of employer learning and to explore whether statistical discrimination is an important source of racial wage gaps.

Table 2: Employer Learning and Statistical Discrimination

<table>
<thead>
<tr>
<th></th>
<th>No Racial Statistical Discrimination</th>
<th>Racial Statistical Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$0 = \beta_{AFQT}^X \leq \beta_{AFQT}^T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$0 &lt; \beta_{AFQT,X}^T = \beta_{AFQT,X}^T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$0 = \beta_{Black}^X \geq \beta_{Black}^T$</td>
<td>$\beta_{Black}^X \leq \beta_{Black}^T &lt; 0$</td>
</tr>
<tr>
<td></td>
<td>$\beta_{Black,X}^T = \beta_{Black,X}^T &lt; 0$</td>
<td>$\beta_{Black,X}^T = \beta_{Black,X}^T &gt; 0$</td>
</tr>
<tr>
<td>Asymmetric Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_{AFQT}^X = \beta_{AFQT}^T = 0$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$0 \leq \beta_{AFQT,X}^T &lt; \beta_{AFQT,X}^T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_{Black}^X = \beta_{Black}^T = 0$</td>
<td>$\beta_{Black}^X = \beta_{Black}^T &lt; 0$</td>
</tr>
<tr>
<td></td>
<td>$0 \geq \beta_{Black,X}^T &gt; \beta_{Black,X}^T$</td>
<td>$0 \leq \beta_{Black,X}^T &lt; \beta_{Black,X}^T$</td>
</tr>
</tbody>
</table>
4 Empirical Results

In this section, we begin the empirical analysis with the replication of the results reported in Altonji and Pierret (2001) using our sample selection criteria. The estimating equation is equation 26 where the time measure is potential experience, and the results are presented in Table 3. The sample in column (1) includes observations coming from interview years 1979-1992, which is the sample used in Altonji and Pierret (2001). Column (2) reports analogous results for our longer sample, the 1979-2008 waves of NLSY79. Because of several differences in sample construction, the results we obtain differ slightly from those presented in Altonji and Pierret (2001). However, their main qualitative results are still presented in Table 3.

Altonji and Pierret (2001) find that employers gradually learn about worker’s productivity as they accumulate more information, but little evidence for statistically discriminate on the basis of race. The results shown in column (1) are supportive for their findings. The coefficients of 0.035(0.014) on \( AFQT \) and 0.068(0.018) on the \( AFQT \)-experience interaction imply that the impact of a one-standard-deviation increase in \( AFQT \) will rise from 0.035 when potential experience is 0 to 0.103 when potential experience is 10 years. Employers learn worker’s productivity from newly acquired information over time, so the weight they put on hard-to-observe correlates of productivity increases over time. If employers obey the law and do not use race as information, then race works as a hard-to-observe correlate of productivity. The initial coefficient on \( Black \) is \(-0.030(0.026)\) and the coefficient on the interaction between \( Black \) and experience is \(-0.084(0.031)\), so the racial wage gap is initially not statistically different from zero, but rising sharply with experience. The regression
Table 3: The Effects of AFQT and Black on Log Wages

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>0.088***</td>
<td>0.071***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Education × Experience/120</td>
<td>-0.034***</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Standardized AFQT</td>
<td>0.035*</td>
<td>0.058***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>AFQT × Experience/120</td>
<td>0.068***</td>
<td>0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.030</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.084**</td>
<td>-0.053***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.274</td>
<td>0.346</td>
</tr>
<tr>
<td>Sample Years</td>
<td>Years 1979-1992</td>
<td>Years 1979-2008</td>
</tr>
<tr>
<td>No.Observations</td>
<td>177288</td>
<td>317988</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of potential experience. All specifications control for urban residence, region of residence, a cubic in experience and years effects. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
coefficients on Black confirms the Proposition 1, therefore, employers do not make use of racial information to determine wage at the time of initial hire.

We obtain qualitative similar results using the longer sample from interview years 1979-2008. However, AFQT and Black now have a flatter profile with experience. Compared with the results in column (1), the absolute value of the coefficients on AFQT and Black are greater, and the coefficients on their interaction terms are much smaller in size, but all statistically significant coefficients in column (1) remain significant. The change in the time paths of AFQT and Black is possibly driven by the non-linear employer learning process. In this article, we assume that the effects of AFQT and Black on log wages vary linearly with time measure to keep the interpretation of coefficients simple.

4.1 The Full Sample

In this section, we test the main predictions of the learning-based statistical discrimination model for our full sample. Our test strategy has two main focal points. First, we analyze how the racial wage gaps vary over time to examine whether or not employers statistically discriminate against black workers. Second, we compare the differences between coefficients of interests in models using experience and using tenure as time measure to test two learning hypotheses.

There is a large body of empirical research on racial wage gap, especially for male workers. It is widely acknowledged that there is a notable black-white wage gap for male workers. In a recent study, Lang and Manove (2011) argue that the substantial racial wage gaps could not be explained by the differences in the quality of schools attended by blacks and by whites, providing evidence that statistical discrimination
is one source of the black-white wage gaps. In that case, our learning-based racial statistical discrimination predicts a narrowing racial wage gap over time because the accumulation of rich information enables employers to base their payments more on worker’s true productivity and less on racial information. If racial statistical discrimination is absent in the labor market, then our model indicates that the racial wage gap will widen resulting from the negative correlation between productivity and race. Therefore, we will examine the time path of racial wage gap to test for statistical discrimination on the basis of race.

To distinguish between symmetric and asymmetric learning, experience and tenure are used as time measure in the wage regression, respectively. From this point on, actual experience will be used as the measure of experience since it is a more accurate measure of worker’s labor market experience and the constructions of actual experience and tenure are more consistent with each other. Symmetric and asymmetric employer learning have different predictions about the initial coefficients on AFQT and Black as well as the size of the coefficients on the interaction terms between models using experience and tenure as time measure. Different initial coefficients and relatively similar coefficients on the interaction terms are supportive for symmetric learning while asymmetric learning implies similar initial coefficients and different coefficients on variables interacted with time measure.

Specifications (1) and (2) in Table 4 estimate the wage equation for the full sample. The coefficients on AFQT-experience and AFQT-tenure interaction terms are both positive and significant, 0.034(0.010) and 0.049(0.017), so we confirm the findings reported in previous studies, that is, employer learning does occur in the labor market. As the worker gains more experience, the employers learn about the
Table 4: The Effects of AFQT and Black on Log Wages by Experience and Tenure

<table>
<thead>
<tr>
<th>Model:</th>
<th>Experience</th>
<th>Tenure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized AFQT</td>
<td>0.068***</td>
<td>0.076***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>AFQT × Experience/120</td>
<td>0.034***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>AFQT × Tenure/120</td>
<td></td>
<td>0.049**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.025</td>
<td>-0.077***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.048**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Black × Tenure/120</td>
<td></td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.033)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.371</td>
<td>0.381</td>
</tr>
<tr>
<td>No. Observations</td>
<td>317988</td>
<td>317988</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of actual experience. Specifications (1) and (2) control for education, education interacted with experience or tenure, urban residence, region of residence, a cubic in experience and years effects. Specification (2) also controls for a cubic in tenure. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
productivity, so the weight on correlate of productivity, $AFQT$, rises over time. The coefficient on $AFQT$ is large in both specifications, implying that employers might directly observe parts of worker’s productivity at the time of initial hire and learn more as additional information become available over time.  

As shown in column (1), the coefficient of $-0.025(0.020)$ on $Black$ is insignificantly negative, and the coefficient of $-0.048(0.018)$ on $Black$ interacted with experience is a significant negative number. The fact that the racial wage gap is initially very small but increasing sharply with experience is in line with the model of no statistical discrimination on the basis of race. A similar finding is reported in Altonji and Pierret (2001) and Mansour (2012).  

The results using tenure as time measure are presented in column (2). At the time of initial entry into the labor market, black workers earn wages that about 7.7 percent lower than those received by their white counterparts with the same AFQT score, and the wage gap decreases insignificant with worker’s tenure, providing some evidence for racial statistical discrimination. Therefore, the use of different time measures in wage regression gives conflicting results, and we can not draw any definite conclusion regarding the issue of racial statistical discrimination. 

\footnote{The assumption of linear relationship between log wage and correlates of productivity may also contribute to the great initial coefficients.}

\footnote{Altonji and Pierret (2001) find little evidence for statistical discrimination in wages on the basis of race, and argue that statistical discrimination plays a relatively unimportant role in the racial wage gap. Mansour (2012) confirms Altonji and Pierret (2001)’s finding of no evidence for racial statistical discrimination, but his empirical results imply that the pattern might differ across occupations. Both studies examine racial statistical discrimination under the assumption of symmetric employer learning.}
To differentiate between two types of employer learning, we need to compare the main coefficients of interest in column (1) and column (2). We test whether the coefficients obtained from these two specifications are significantly different from each other. The $P$-values for the difference between these two sets of coefficients indicates that the coefficients on $AFQT$ are not significantly different from each other (0.291), but the coefficient on Black-tenure interaction term differs significantly from that on Black-experience interaction term (0.017), providing evidence in favor of asymmetric learning. However, the significantly different coefficients on $Black$ (0.001) and qualitatively similar coefficients on $AFQT$ interacted with experience and tenure (0.323) are consistent with the predictions of symmetric learning. Overall, the empirical evidence on whether employer learning is symmetric or asymmetric is inconclusive. 17

To summarize, the full sample gives confusing results concerning employer learning and statistical discrimination. The results using experience as time measure implies the absence of racial statistical discrimination but the use of tenure measure leads to an opposite conclusion. The comparison between the results using experience and using tenure also provides mixed support for the nature of employer learning.

17 The existing empirical literature offers no conclusive evidence on the nature of employer learning. Schonberg (2007) find that employer learning is mostly symmetric even though there are important differences across educational groups. Pinkston (2009)’s empirical results suggest that asymmetric employer learning plays a role that is at least as much important as symmetric learning during an employment spell.
4.2 High School Graduates and College Graduates

An important finding in the employer learning literature is that the employer learning process varies across different educational groups. Arcidiacono, Bayer, and Hizmo (2010) propose that employers gradually learn about the productivity of high school graduate, but for college graduates, employers nearly perfectly observe their productivity at the time of hire. They conclude that a college degree helps workers directly reveal key aspects of productivity, and thus employer learning is more important for high school graduates. Schonberg (2007) also finds that there are important differences across educational groups with respect to the process of employer learning. When workers involved are college graduates, the empirical evidence is potentially consistent with asymmetric learning model.

In this section, we focus our attention on high school graduates and college graduates, and perform the empirical analysis for these two educational groups. We estimate the wage equations 27 and 28 on these two educational groups, respectively, and the education variable and its interaction terms are omitted from the regressions since the education level does not vary much among each group.

Table 5 presents the estimating results for high school graduates and college graduates. Specifications (1) and (2) in Table 5 estimate the wage regression for the high school sample where actual experience and tenure are used as measure of time, respectively. The coefficients on AFQT interaction terms are positive and statistically significant, 0.035(0.011) and 0.075(0.023), suggesting that employers learn about the productivity of high school graduates gradually over time. The significantly negative coefficient on Black in both specifications (1) and (2), −0.057(0.027) and −0.099(0.024), provides empirical evidence for statistical discrimination on the
Table 5: The Effects of AFQT and Black on Log Wages by Experience and Tenure by Education

<table>
<thead>
<tr>
<th>Model:</th>
<th>High School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Standardized AFQT</td>
<td>0.060***</td>
<td>0.062***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>AFQT × Experience/120</td>
<td>0.035**</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>AFQT × Tenure/120</td>
<td>0.075**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.057*</td>
<td>-0.099***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Black × Tenure/120</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.219</td>
<td>0.235</td>
</tr>
<tr>
<td>No.Observations</td>
<td>153456</td>
<td>153456</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of actual experience. All specifications control for urban residence, region of residence, a cubic in experience and years effects. Specifications (2) and (4) also control for a cubic in tenure. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* p < 0.05, ** p < 0.01, *** p < 0.001
basis of race for high school graduates. Employers know the average productivity of black workers are lower than that of white workers, and use race as information to determinate wage for newly hired workers. They pay black workers less than apparently similar white workers because black workers are associated with lower average productivity. The coefficient of $-0.023(0.022)$ on Black interacted with experience and the coefficient of $0.059(0.047)$ on Black interacted with tenure suggest that the racial wage gap does not change much over experience but decreases insignificantly over tenure. Consistent with the predictions of racial statistical discrimination, employers reduce their reliance on racial information as additional information about worker’s productivity become available, so the racial wage gap declines over time.

If employer learning is asymmetric, the use of different time measures, experience or tenure, in the wage regression will yield qualitatively similar coefficients on $AFQT$ and $Black$, but significantly different coefficients on these variables interacted with time measure. The coefficients on $AFQT$ are quite similar, $0.060(0.014)$ using experience and $0.062(0.013)$ using tenure, and a $P$-value of 0.831 fails to reject the null hypothesis that these two coefficients are equal. In contrast, the coefficient on $AFQT$ interacted with experience, $0.035(0.011)$, is much smaller that on AFQT interacted with tenure, $0.075(0.023)$, and these two coefficients are significantly different from each other. The quantitatively similar coefficients on $AFQT$ and significantly different coefficients on its interaction terms indicate that employer learning is more likely to be asymmetric. As discussed earlier, the coefficients on $Black$ using either experience or tenure are significantly negative, $-0.057(0.027)$ and $-0.099(0.024)$. However, the equality of these two coefficients are rejected at the significant level of 5%, conflicting the predictions of asymmetric learning. The comparison between the
The coefficient on Black interacted with experience and that on Black interacted with tenure provides further empirical evidence supporting asymmetric learning. The racial wage gap only slightly changes with experience but decreases greatly over tenure, and these two coefficients significantly different from each other.

Overall, the results for high school graduates presented in column (1) and (2) of Table 5 are generally supportive for asymmetric learning and statistical discrimination by race. The only exception is that the equality of coefficients on Black using different time measures are rejected by the test even though both coefficients have the predicted negative sign. One possible explanation is that employers statistically discriminate against black workers more on jobs require prior experience than on entry-level jobs.

Columns (3) and (4) of Table 5 show the regression results for college graduates. In both specifications, the coefficient on AFQT is large and statistically significant but the coefficient on AFQT interacted with time measure is insignificant and relatively small. The time trend of the AFQT coefficient provides little evidence for employer learning, therefore, employers have nearly perfect information about the productivity of newly hired college graduates. In contrast to high school graduates, college-educated black workers earn a higher wage than their white counterparts whenever they first enter the labor market or make a job change, and this black wage premium declines over time. This initial wage premium for black college graduates is also reported in Arcidiacono, Bayer, and Hizmo (2010). The existence of a substantial black wage premium for college graduates actually is a robust feature of the U.S. labor market.

The results for college graduates shown in Table 5 confirm the previous findings
presented in Arcidiacono, Bayer, and Hizmo (2010) that the key aspects of productivity are directly revealed upon initial entry into the college market. Employers almost perfectly observe the productivity of college graduates at the time of initial hire, and learn very little additional productivity over time. They explain 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. Therefore, the learning-based racial statistical discrimination model developed in this article is not applicable to the group of college graduates.

A related finding from the empirical research on racial wage gap is that the wage gap between blacks and whites is smaller or even non-exist for high-skilled workers. Neal and Johnson (1996) claims that the black-white wage gap for male declines with the skill level, and a similar finding is also reported in Lang and Manove (2011), black and white men have similar earnings at high levels of education and AFQT score. Our learning-based racial statistical discrimination provides a plausible explanation for the lack of racial wage gap among high-skilled workers. In high-skill labor market primarily dominated by a better-educated workforce, employers have less incentives to statistically discriminate against black workers because they could accurately assess worker’s productivity.

The results for the full sample shown in Table 4 are the mixed results of different educational groups. The reason why we are unable to obtain clear-cut results for the full sample is that the employer learning process varies across different educational groups. We focus on high school graduates and college graduates in our empirical analysis since the sample sizes for high school and college drop-outs are smaller and their results are much less statistically significant.
5 Robustness Checks

5.1 Pure or Imperfect Asymmetric Learning

Our learning-based racial statistical discrimination model assumes that learning is either purely symmetric or purely asymmetric. Pure symmetric learning means that current and outside firms have the same information about worker’s productivity. Pure asymmetric learning, in contrast, suggests that only current firms accumulate information about worker’s productivity and the outside firms receive no new information.

It is likely that outside firms receive some new information about worker’s productivity but current firms have superior information than outside firms do, then employer learning is termed as imperfectly asymmetric. In the case of imperfect asymmetric learning, outside firms could learn some aspects of worker’s productivity, so employer learning occurs both over experience and over tenure. Schonberg (2007) analyses how the impacts of ability and education vary with experience and tenure to distinguish between pure symmetric and imperfect asymmetric learning. If learning is purely symmetric, then the coefficients on education interacted with tenure and AFQT interacted with tenure should both be zero since these variables interacted with experience are included in the regression model. In contrast, non-zero coefficients on tenure interaction terms suggest that current firms have information advantage over outside firms. Her empirical results using the same NLSY79 data provide evidence for imperfect asymmetric learning among college graduates.\(^\text{18}\)

\(^\text{18}\)Schonberg (2007) uses the 1979-2001 waves of NLSY79 and restricts the empirical analysis to white males only.
We apply a methodology similar to Schonberg (2007)’s to examine the possibility that employer learning is imperfectly asymmetric instead of purely asymmetric. Our empirical results imply that employer learning process mainly occurs in the high school market, so we only focus on high school graduates here. If some new productivity information is revealed to outside firms, then the learning process takes place not only over tenure but also over general experience, suggesting that employer learning is imperfectly asymmetric. On the other hand, if outside firms are completely excluded from the employer learning process, then employer learning is purely asymmetric and we should only observe learning over tenure. We include both experience and tenure interactions in the regression model and our estimating equation is given by

\[ \ln w_{i,t} = \beta_0 + \beta_1 AFQT_i + \beta_2 (AFQT_i \times T_{i,t}) + \beta_3 (AFQT_i \times X_{i,t}) + \beta_4 Black_i + \beta_5 (Black_i \times T_{i,t}) + \beta_6 (Black_i \times X_{i,t}) + \beta_7 \Omega_{i,t} + \epsilon_{i,t}. \] (29)

The main coefficients of interest are \( \beta_3 \), the coefficient on the interaction between AFQT and experience, and \( \beta_6 \), the coefficient on the interaction between Black and experience. Pure asymmetric learning predicts that both \( \beta_3 \) and \( \beta_6 \) should be equal to zero while imperfect asymmetric learning indicates non-zero coefficients on experience interaction terms.

The estimating results under the assumption of pure asymmetric learning presented in column (2) of Table 5 appear again in column (1) of Table 6, for the sake of comparison, and column (2) shows the empirical results applying regression equation 29 where both tenure and experience interactions are considered. The fact that
### Table 6: Pure or Imperfect Asymmetric Learning

<table>
<thead>
<tr>
<th>Model:</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized AFQT</td>
<td>0.062***</td>
<td>0.053***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>AFQT × Tenure/120</td>
<td>0.075**</td>
<td>0.062*</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>AFQT × Experience/120</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.099***</td>
<td>-0.073**</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Black × Tenure/120</td>
<td>0.059</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.235</td>
<td>0.236</td>
</tr>
<tr>
<td>No. Observations</td>
<td>153456</td>
<td>153456</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of actual experience. Specifications (1) and (2) control for urban residence, region of residence, a cubic in experience, a cubic in tenure, and years effects. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \)
the coefficients on AFQT-experience and black-experience are not statistically different from zero provides empirical evidence in favor of pure asymmetric learning. Unlike the experience interactions, the coefficients on tenure interactions are big in magnitude and statistically significant.\(^{19}\) To further distinguish between pure and imperfect asymmetric learning, we test whether the common regression coefficients presented in specification (1) and (2) are significantly different from each other. The \(P\)-values suggest that there are no statistical differences between these two sets of coefficients, therefore, the employer learning is more likely to be purely asymmetric. The empirical results presented in Table 6 indicate that outside firms seem not to have access to new information about worker’s productivity accumulated over time, implying that, for high school graduates, employer learning tends to be more supportive of purely asymmetric rather than imperfectly asymmetric.

### 5.2 Match Quality and Job Mobility

In order to concentrate on employer learning and statistical discrimination, the theoretical model developed in this article does not consider the quality of firm-worker match. If various racial groups are associated with different job mobility patterns, the match quality might vary between black and white workers.

Our empirical results support the view that, in the high school market, employers use race as a source of information to infer worker’s productivity and thus statistically discriminate against black workers. Whenever workers start a new job, employers pay black high school graduates significantly less than their white counterparts con-

\(^{19}\)The coefficient of 0.090(0.056) on black interacted with tenure is slightly significant at the 10% significant level.
ditional on the AFQT score. One possible explanation for the racial wage gap is that black high school graduates on average have a worse firm-worker match quality than do white high school graduates. If the racial wage gap reflects racial differences in match quality rather than statistical discrimination, we could expect that black high school graduates generally switch firms more frequently than do white high school graduates. Poorly matched black workers are more likely to move between firms voluntarily or involuntarily.

We test this alternative explanation by estimating a probit model that examines the effect of black indicator on the worker's probability of job change. We apply the following estimating equation to high school graduates and college graduates separately:

\[
P_r(\text{JobChange}_{i,t} = 1) = \Phi(\beta_0 + \beta_1 Black_i + \beta_2 AFQT_i + \beta_3 X_{i,t} + \beta_4 \Omega_{i,t}),
\]

where the main coefficient of interest is \( \beta_1 \). If the racial wage gap in the high school market could be attributed to match quality differences across racial groups instead of racial statistical discrimination that resulting from employer learning, we could expect a positive \( \beta_1 \) in the high school sample. That is, black high school graduates change job more frequently than their white counterparts.

The results of the probit regression are presented in Table 7. As shown in column (1), the probability of job change is 3.6% lower for black workers in the high school market. The marginal effect of the black indicator is statistically significant at the 5% significant level. The fact that black workers are less likely to change job rules out the alternative interpretation that the racial wage gap in the high school market reflects racial differences in firm-worker match quality. The results for college graduates appear in column (2). It turns out that there is no statistical racial difference in
Table 7: The Impact of the Black Indicator on the Probability of Job Change

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>-0.036*</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Standardized AFQT</td>
<td>-0.022*</td>
<td>-0.074***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Sample</td>
<td>High School</td>
<td>College</td>
</tr>
<tr>
<td>No. Observations</td>
<td>466704</td>
<td>181452</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is a dummy variable for job change. Specifications (1) and (2) control for actual experience, urban residence, region of residence and years effects. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
the probability of job change among college graduates. The probit regression results from the college graduates sample provide further empirical evidence for the learning-based racial statistical discrimination model established in this article. In the high school market, black workers face statistical discrimination at the time of initial hire, therefore, black workers tend to change job less frequently than do white workers to mitigate the effects of discrimination. In the college market, employers could directly observe key aspects of worker’s productivity and learn very little additional information over time, so employers have less incentives to discriminate against black workers. As a consequence, the job change rate does not vary between black and white college graduates.

The results shown in Table 7 also indicates that workers of high productivity might have better match quality than low-ability workers. A one standard deviation increase in the AFQT score reduces the job change rate of high school graduates and college graduates by 2.2% and 7.4%, respectively. The positive association between productivity and match quality is stronger in the college market than in the high school market. The negative marginal effect of the AFQT score on the probability of job change obtained from the high school graduates sample is also consistent with one well-known consequence of asymmetric learning, the adverse selection of job movers. The productivity of workers who switch firms is lower than that of workers who stay with their employers.\(^\text{20}\)

\(^{20}\text{Gibbons and Katz (1991) argue that laid-off workers are generally less able than exogenous movers. Their empirical results that the wage loss for laid-off workers is greater than that for exogenous movers support this view. In this article, however, we do not distinguish between involuntary and voluntary job movers due to data limitations.}\)
In short, the results of the probit model provide several pieces of evidence in favor of our main findings. First, the fact that black high school graduates change job less frequently than do white high school graduates strengthens our previous finding that black workers are statistically discriminated in the high school market. Furthermore, there is no statistical racial difference in job mobility patterns in the college market, confirming that employers almost perfectly observe the productivity of college-educated workers and therefore have weak incentives to discriminate against blacks. Second, the negative and statistically significant marginal effect of AFQT score on job change rate are consistent with the previous finding that employer learning is asymmetric in the high school market.

5.3 Occupation and Industry

Our employer learning framework does not consider the role of job assignment. It is likely, however, that black and white workers are associated with jobs of different skill-level. If that is the case, there may be explanations other than learning-based racial statistical discrimination for our findings for high school graduates. One possible alternative explanation is that black workers are more likely to be hired into low-skill-level jobs at the start of their career and to be trapped for a while in such jobs.21 What appears to be evidence of racial statistical discrimination might could be attributed to differences in the skill level of jobs taken by black and white workers.

To test the possibility that racial wage gap is driven by blacks and whites being sorted into jobs of different skill-level, we add the initial occupation variable to the estimating equation as an additional control, and repeat the empirical analysis.

21The initial job assignments could influence the entire menu of career paths.
separately for high school graduates and college graduates. The regression results are presented in Table 8.

Table 8: The Effects of Black on Log Wages Controlling for Initial Occupation

<table>
<thead>
<tr>
<th>Model:</th>
<th>High School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.084***</td>
<td>-0.122***</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.014</td>
<td>-0.103**</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Black × Tenure/120</td>
<td>0.067</td>
<td>-0.206**</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.242</td>
<td>0.255</td>
</tr>
<tr>
<td>No. Observations</td>
<td>119496</td>
<td>119496</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of actual experience. All specifications control for urban residence, region of residence, a cubic in experience, years effects and initial occupation. Specifications (2) and (4) also control for a cubic in tenure. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

* \(p < 0.05\), ** \(p < 0.01\), *** \(p < 0.001\)

The empirical results suggest that occupation sorting could not explain the results that I attribute to learning-based racial statistical discrimination. We obtain qualitatively similar results with the inclusion of initial occupation, so our main find-

\footnote{We distinguish 7 occupations: professional workers; managers; sales workers; clerical workers; craftsman and operatives; agricultural labors; and service workers.}
ings still hold. In the high school market, the Black coefficient is still significantly negative, and rises insignificantly with tenure, providing evidence that our main results can not be attributed to differences in the skill-level of jobs taken by different racial groups. Employers initially statistically discriminate against black high school graduates but learn about their productivity gradually as more information become available, so the racial wage gaps shrink over time. Including the initial occupation in the regression also does not alter the results for college graduates. College-educated blacks do not face statistical discrimination on the basis of race, instead, they earn an initial black wage premium conditional on productivity.

We further explore the issue of job assignment by examining the role of initial industry in the observed racial wage gap. With initial industry added, we redo the empirical analysis for two educational groups. The results shown in Table 9 closely resemble those of Table 8. Therefore, the racial wage gap can not be explained by the possible variation in industry that members from different racial groups work in.

\[23\text{Mansour (2012)} \text{ finds that there is substantial variation in the time path of black coefficients across occupations. The results shown in Table 8 indicate that the inclusion of initial occupation does affect the coefficients on Black and black interacted with time measure. Therefore, the extent of statistical discrimination on the basis of race varies across occupations.}\]

\[24\text{We distinguish 12 industries: agriculture; mining; construction; manufacturing; transportation, communication, and utilities; wholesale and retail trade; finance, insurance, and real estate; business and repair services; personnel services; entertainment and recreation services; professional and related services; and public administration.}\]

\[25\text{We also experiment by controlling for initial occupation and industry simultaneously, the main results are not affected. The empirical results are available upon request.}\]
Table 9: The Effects of Black on Log Wages Controlling for Initial Industry

<table>
<thead>
<tr>
<th>Model:</th>
<th>High School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.096**</td>
<td>-0.136***</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Black × Experience/120</td>
<td>-0.017</td>
<td>-0.089*</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Black × Tenure/120</td>
<td>0.066</td>
<td>-0.185*</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.251</td>
<td>0.263</td>
</tr>
<tr>
<td>No.Observations</td>
<td>119124</td>
<td>119124</td>
</tr>
</tbody>
</table>

Notes: The experience measure is years of actual experience. All specifications control for urban residence, region of residence, a cubic in experience, years effects and initial industry. Specifications (2) and (4) also control for a cubic in tenure. The numbers in parentheses are White/Huber standard errors accounting for multiple observations per person.

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

In this article, we combine elements of both employer learning theory and screening discrimination theory to develop a learning-based racial statistical discrimination model. We formulate a framework that nests both types of employer learning, and examine whether employers statistically discriminate against black workers without the assumption of symmetric learning.

Our estimating results show that high school 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 high school 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 after hire. As a result, we mainly focus our attention on high school graduates. The time paths of racial wage gap in the high school market indicate that employers use race as information to infer worker’s productivity and black workers are statistically discriminated. By comparing the coefficients of interests from wage regressions using general experience and job tenure as time measure, we find empirical evidence for asymmetric learning in the high school market. As 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 are unable to update their expectations about worker’s productivity.

In this article, we simply assume that the effects of AFQT score and racial vari-
able on log wage vary linearly with time when performing the empirical analysis. However, the theoretical model predicts a varying speed of employer learning. **Lange (2007)** constructs a framework to formally estimate the speed of employer learning and shows that employers actually learn very fast.  

One direction for future research is to precisely estimate the speed of employer learning and to test one important implication of our model, that is, employers learn faster over tenure than over experience in the case of asymmetric learning. The speed of employer learning is also crucial for the economic significance of statistical discrimination on the basis of race. The faster the employers learn, the shorter the time period during which employers need to rely on race to predict a worker’s productivity.

Most screening discrimination models build around the assumption that the signal of productivity employers receive is less reliable for black workers than for white workers at the time of initial hire. 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 know the average productivity of black workers is lower than that of white workers. **Pinkston (2006)** applies the framework of employer learning to test the hypothesis that the signal employers receive from black workers is nosier than that of white workers. His estimation results provide evidence supporting this view. 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.

---

26 **Lange (2007)** shows that the initial expectation errors decline by 50% within 3 years.
References


