

# Statistical Discrimination, Employer Learning, and Employment Gap by Race and Education\*

Seik Kim\*\* · Hwa Ryung Lee\*\*\*

*Tests of statistical discrimination require evaluation records provided by employers or variables that employers do not observe directly but are observed by researchers. As such variables are difficult to obtain, this paper develops a strategy that uses variables available in usual data sets. This paper derives testable implications for statistical discrimination by exploiting the heterogeneity in employer learning processes. Evidence from analysis using the March Current Population Survey for 1971-2016 is consistent with the theoretical predictions. The empirical findings are not explained by alternative hypotheses, such as human capital theory, taste-based discrimination, or search and matching models.*

JEL Classification: D83, J64, J71

Keywords: Employer Learning, Statistical Discrimination, Unemployment Rate

## I. Introduction

In hiring and wage-setting processes, employers make judgments about the value of workers using all the information available at the time of their decision-making. The productivity of workers, however, is never perfectly observed, and employers must make predictions on the basis of limited information. At the time of labor

---

*Received: March 22, 2018. Revised: May 18, 2018. Accepted: June 12, 2018.*

\* The authors have benefited from helpful comments made by Joseph Altonji, Yoram Barzel, Byung-hill Jun, Fabian Lange, Shelly Lundberg, two anonymous reviewers, and seminar participants at the annual meetings of the Society of Labor Economists, European Association of Labor Economists, Population Association of America, European Society for Population Economics, Society for the Study of Economic Inequality, Academia Sinica, Korea Institute of Public Finance, Korea University, Sogang University, University of Zurich, University of Washington (UW), and the UW West Coast Poverty Center (WCPC). The authors gratefully acknowledge support from the WCPC through the Emerging Poverty Scholar Small Grant. This work was supported by a Korea University Grant (K1509061).

\*\* First Author, Associate Professor, Department of Economics, Korea University, 145 Anam-Ro, Seongbuk-Gu, Seoul 02841, Korea. Phone: 82-2-3290-5135. E-mail: seikkim@korea.ac.kr.

\*\*\* Corresponding Author, Fellow, Korea Development Institute, 263 Namsejong-Ro, Sejong 30149, Korea. Phone: 82-44-550-4101. E-mail: hwaryung.lee@kdi.re.kr.

market entry, for example, potential workers do not have past labor market experience, and employers receive only noisy signals on worker productivity, such as curriculum vitae, recommendation letters, and interview results, as well as easily observable indicators such as race or education. Perceived productivity is likely to differ from actual productivity. In particular, if employers are less able to evaluate the productivity of one worker relative to another, even equally productive workers may face unequal opportunities in the labor market. This outcome is referred to as screening discrimination by Cornell and Welch (1996), and is a type of statistical discrimination.<sup>1</sup>

As young workers gain more experience, their past labor market performance records become available to employers, who are then able to make better predictions about the workers' future performance. The theory of statistical discrimination, accompanied by the employer learning hypothesis, predicts that the degree of discrimination would decrease with the labor market experience of workers. Altonji and Pierret (2001) utilize this idea and propose an empirical test for statistical discrimination. Consider the variables that are correlated with productivity. Some are observable by both employers and researchers (e.g., race or education), while others are observed by researchers only (e.g., test scores). Using the 1979 National Longitudinal Study of Youth (NLSY79), they show that if employers statistically discriminate among young workers on the basis of easily observable characteristics, the coefficients of the easily observed variables in a wage equation should fall and the coefficients of hard-to-observe variables should rise as the workers gain experience.

Tests of statistical discrimination require variables available to researchers but not observed by employers (Altonji and Pierret, 2001; Pinkston, 2006) or data on employer-provided performance measures (Neumark, 1999; Pinkston, 2003), but such variables are difficult to obtain in practice. To overcome the data availability problem, this paper develops an identification strategy that does not rely on those specific variables.<sup>2</sup> The data requirement for the proposed strategy is minimal. Suppose that employers statistically discriminate among young workers on the basis of easily observable characteristics such as race or education, but learn about their productivity over time. This paper shows that the unemployment rates for disadvantaged groups will be higher than for non-disadvantaged groups at the time

---

<sup>1</sup> The idea that the precision of productivity signals across workers from different groups is a source of statistical discrimination dates back to the seminal paper by Phelps (1972).

<sup>2</sup> A test proposed by Oettinger (1996) also does not require such variables, but requires job mobility and job tenure information. His model suggests that the gain from job changes for African-American men should be smaller than that for white men. As a result, African-American men should move less and the black-white difference in wages among men should increase with experience. Also not using those variables, Moro (2003) structurally estimates an equilibrium labor market model with statistical discrimination. His model, however, does not prove statistical discrimination as the results can also be generated by taste-based discrimination.

of labor market entry and will decline faster with experience. Therefore, statistical discrimination in the presence of employer learning can be detected by individual-level repeated cross-section data on employment status, experience, and the variables on which discrimination is based, such as race or education.

This paper focuses on employment opportunities rather than wage levels because discrimination will influence the former more than the latter if the Equal Employment Opportunity Act prohibits wage differences among workers performing the same task. An obstacle to using this approach, however, is that employment status and wages provide different degrees of information: employment is measured as a binary variable, whereas wages are measured continuously. Moreover, as this study does not use panel data nor the specific variables employed in the previous literature, testing statistical discrimination requires more steps compared with existing methods.<sup>3</sup> For example, to show that the theoretical predictions hold empirically, the results should not be explained by other hypotheses, such as human capital theory, taste-based discrimination, or search and matching models. This paper demonstrates that the predictions of these alternative hypotheses are not consistent with the empirical findings.

This study is specifically interested in statistical discrimination on the basis of race and education as the noise in productivity signals is likely to depend on both race and education. In addition, both race and education have their own distinctive features. For example, education is a determinant of productivity, but race is not. The Title VII of the Civil Rights Act prohibits employment discrimination based on race as well as color, religion, sex, and national origin, but not on education. Although the act lists a number of items which must not be used as a device for discrimination, in this research race is of particular interest. Color and religion are of less interest in this research as they are not often available in the data. Gender discrimination is not covered as women participating in the labor force are not randomly selected. Discrimination on the basis of national origin is excluded because of the difficulty of separating employer learning from economic assimilation.

This paper proceeds as follows. Section 2 presents theoretical predictions about workers' employment opportunities over their career when employers statistically discriminate and update their predictions using workers' labor market history. It produces three testable implications. Section 3 applies the proposed strategy using the March Current Population Survey (CPS) for 1971–2016. It discusses descriptive statistics, analyzes estimation results, and shows that alternative explanations other than statistical discrimination are likely to be rejected. Section 4 conducts several

---

<sup>3</sup> Ritter and Taylor (2011) use the NLSY79 to examine whether the black-white employment gap can be explained by the associated disparity in AFQT scores, a variable available to researchers but not observed by employers. They find a large unexplained unemployment differential and explain it using a model that utilizes statistical discrimination.

robustness checks to verify the main findings of this research. Section 5 concludes.

## II. Theoretical Framework

### 2.1. Worker Productivity

Suppose that potential worker  $i$  enters the labor market for the first time. He is characterized by productivity,  $P_{ij}$ , when he is matched with employer  $j$ . The productivity depends on two sets of human capital measures,  $H_i$  and  $\eta_i$ . Vector  $H_i$  consists of variables that are directly observed by both employers and researchers. These variables include years of schooling, labor market experience, and possibly job tenure.<sup>4</sup> In general, race is not an element of vector  $H_i$  as it does not affect productivity directly once other observable human capital measures are controlled for.<sup>5</sup> Vector  $\eta_i$  consists of a finite number of skill measures. Examples of these skills include major in college, academic performance, physical strength, and communicative skills. Although these measures are generally unobservable by researchers, they may be partly observed by some employers. We assume that  $\eta_i$  has a multivariate normal distribution.

As different jobs require different skills, worker  $i$ 's productivity in job  $j$  is specified through a linear combination of these factors,

$$P_{ij} = r'H_i + r_j'\eta_i, \quad (1)$$

where  $r$  is a vector of parameters common to all employers and  $r_j$  is an employer-specific non-random weight. This setup of productivity is an extension of Lundberg and Startz (1983) using job-specific weights for a vector of ability measures. The quality of a match depends on  $r_j'\eta_i$ , the sum of worker  $i$ 's skills weighted by employer  $j$ 's skill weight. As the weight vector,  $r_j$ , is employer-specific, the workers who are most productive in one firm would not necessarily be the most productive in another firm.<sup>6</sup>

<sup>4</sup> At the time of labor market entry, worker  $i$  has zero labor market experience and job tenure.

<sup>5</sup> Race may still be correlated with productivity if human capital measures are not controlled for. In this sense, our baseline assumption is in line with the assumptions that have been made in the previous literature. For example, Altonji and Pierret (2001) and Lange (2007) assume that a worker's ability is correlated with group membership to explain why employers use race as a device to predict a worker's unobservable skills at the beginning of his career.

<sup>6</sup> Letting skill weights be employer-specific is a more intuitive way to understand why an employed worker changes jobs and how an unemployed worker gains employment. Allowing  $r_j$  to be common to all  $j$  does not affect the theoretical content of this paper.

## 2.2. Perceived Productivity at the Time of Labor Market Entry

When employer  $j$  makes predictions about applicant  $i$ 's productivity, she has complete information about  $H_i$  and partial knowledge of the skill measures comprising  $\eta_i$ . Formally, let  $I_{ij}$  denote the set of information that employer  $j$  has about applicant  $i$  when person  $i$  enters the labor market.<sup>7</sup> Such an information set includes observable human capital measures in  $H_i$  as well as factors that are not necessarily correlated with productivity such as race. It also includes employer  $j$ 's private information about applicant  $i$ . Therefore, an information set is worker-employer-specific, and different employers, even those with the same  $r_j$ , have different abilities to screen workers. For example, if applicant  $i$  and employer  $j$  share a similar cultural background, but applicant  $i'$  and employer  $j'$  do not,  $I_{ij}$  will be richer than  $I_{i'j}$  or  $I_{ij'}$ , all other things being equal. Cultural background is broadly defined, as in Cornell and Welch (1996), to include groups characterized by language, ethnicity, school ties, neighborhood connections, or membership of social organizations, as well as age, race, sex, and education. Hence, employers  $j$  and  $j'$  may rank the same applicant differently even when  $r_j$  and  $r_{j'}$  are the same. Under this setting, applicant  $i$ 's true productivity (1) is perceived by employer  $j$  as

$$E[P_{ij} | I_{ij}] = r'H_i + r'_j E[\eta_i | I_{ij}]. \quad (2)$$

From employer  $j$ 's point of view, (2) implies that  $r'H_i$  is deterministic and  $r'_j\eta_i$  is partially observable.<sup>8</sup> Whether  $E[\eta_i | I_{ij}]$  in (2) is a good predictor of  $\eta_i$  depends on employer  $j$ 's private information,  $I_{ij}$ . If employer  $j$  has perfect knowledge about applicant  $i$ 's productivity, then the perceived productivity will be equal to the actual productivity,  $E[\eta_i | I_{ij}] = \eta_i$ . If no information is available, employer  $j$ 's best prediction will be simply the expected value,  $E[\eta_i | I_{ij}] = E[\eta_i]$ .

While  $I_{ij}$  is worker-employer-specific, employers may evaluate the productivity of one group of potential workers systematically better than that of another. Specifically, suppose that employers can easily categorize potential workers into two groups, group A and group B. Without loss of generality, assume that employers have richer information sets for group A workers than for group B workers,

$$I_{i \in A, j} \supset I_{i \in B, j} \quad \text{and} \quad I_{i \in A, j} \neq I_{i \in B, j} \quad \text{for any } j. \quad (3)$$

Condition (3) is equivalent to assuming that, for any employer, the signals from

<sup>7</sup> Therefore, the distribution of perceived productivity differs from the distribution of actual productivity. See, for example, Cho, Chu, Kim, and Lee (2016).

<sup>8</sup> In contrast,  $r'_j\eta_i$  is purely stochastic from the researcher's point of view.

group B workers are noisier than those from group A workers. In practice, employers may use a variety of measures to classify workers. We consider two such measures: race and education. In this paper when race is used to identify group membership, group A members are non-Hispanic white workers and group B members are African-American workers. When education is used, group A members are those more educated and group B members are those less educated. Groupings do not have to be dichotomous and can even be continuous, but to keep the discussion simple we stick to group A and group B.

Let  $S_{ij}$  denote the signal that employer  $j$  receives from applicant  $i$ 's  $\eta_i$ . As a representation of (3), consider

$$S_{i \in A, j} = r'_j \eta_i + \xi_{i \in A, j} \quad \text{and} \quad S_{i \in B, j} = r'_j \eta_i + \xi_{i \in B, j}, \quad (4)$$

where  $\xi$  is a normal random variable independent of  $\eta_i$  with  $E[\xi_{i \in A, j}] = E[\xi_{i \in B, j}] = 0$  and  $Var(\xi_{i \in A, j}) \equiv \sigma_{A\xi}^2 < Var(\xi_{i \in B, j}) \equiv \sigma_{B\xi}^2$ . Assumption (4) implies that African-American workers have noisier productivity signals than non-Hispanic white workers or similarly that less educated workers have noisier productivity signals than more educated workers.<sup>9</sup> We use a common subscript  $i$  in (3) and (4) to emphasize that the distribution of the stochastic part of productivity in (1),  $r'_j \eta_i$ , is common to both group A and group B individuals. Using a common subscript  $i$  enables us to prove that statistical discrimination can arise whether there exists the productivity gap between group A and group B workers or not. For example, a low-educated non-Hispanic white worker  $i_1$  and a high-educated non-Hispanic white worker  $i_2$  have different deterministic terms,  $r'H_{i_1}$  and  $r'H_{i_2}$ , and therefore have different wages, but the distributions of productivity less the deterministic term,  $r'_j \eta_i$ , are identical by construction.

The information gap given in (3) has an important implication for the distributions of employers' expectations in (2). As the productivity of a group A member is more precisely observed than that of a group B member, the ex ante variance of employer  $j$ 's perceived productivity of group A members is strictly larger than the ex ante variance of employer  $j$ 's perceived productivity of group B members,

$$Var(E[P_{ij} | I_{i \in A, j}]) > Var(E[P_{ij} | I_{i \in B, j}]) \quad \text{for any } j. \quad (5)$$

To better understand (5), consider an extreme case in which employer  $j$  does

<sup>9</sup> In practice, employers are disproportionately non-Hispanic whites, and therefore African-American workers have noisier productivity signals than non-Hispanic white workers. Similarly, the productivity of college graduates is more precisely observed than that of high school graduates as college graduates have majors.

not have any screening ability for group B members. Then, the employer will evaluate the productivity of any group B worker as the unconditional expectation,  $E[\eta_i | I_{i \in B, j}] = E[\eta_i]$ , and researchers have  $Var(E[P_{ij} | I_{i \in B, j}]) = Var(E[r'_j \eta_i | I_{i \in B, j}]) = 0$ . Another extreme example is the case in which employer  $j$  has perfect knowledge about the productivity of worker  $i$  who is a group A member,  $E[\eta_i | I_{i \in A, j}] = \eta_i$ . In this case, the employer perfectly observes the productivity distribution of group A workers and researchers have  $Var(E[P_{ij} | I_{i \in A, j}]) = Var(E[r'_j \eta_i | I_{i \in A, j}]) = Var(r'_j \eta_i)$ .

An alternative way of deriving (5) is by using the information structure given in (4). That is, (2) can be rewritten as

$$E[P_{ij} | I_{ij}] = E[P_{ij} | S_{ij}, H_i] = r'H_i + E[r'_j \eta_i | r'_j \eta_i + \xi_{ij}] = r'H_i + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} S_{ij}, \quad (6)$$

where  $\sigma_\eta^2 \equiv Var(r'_j \eta_i)$ .<sup>10</sup> In the first case given above, the signal is extremely noisy,  $\sigma_{B\xi}^2 = \infty$ , and all applicants from group B will be evaluated as  $r'H_i$ . In the second case, the signal is perfect,  $\sigma_{A\xi}^2 = 0$ , and employers perfectly observe individual productivity as being  $r'H_i + r'_j \eta_i$ . Using the information structure given in (4), we can verify (5) by

$$Var(E[P_{ij} | I_{i \in A, j}]) = \frac{\sigma_\eta^4}{\sigma_\eta^2 + \sigma_{A\xi}^2} > \frac{\sigma_\eta^4}{\sigma_\eta^2 + \sigma_{B\xi}^2} = Var(E[P_{ij} | I_{i \in B, j}]) \text{ for any } j. \quad (7)$$

since  $\sigma_{A\xi}^2 < \sigma_{B\xi}^2$ . As the information set  $I_{ij}$  gets richer, the distribution of employer  $j$ 's perceived productivity,  $r'_j \eta_i | I_{ij}$ , approaches the distribution of the actual productivity,  $r'_j \eta_i$ .

### 2.3. Initial Job Offers and Unemployment Rates

Since the productivity of a group A worker is more precisely observed than that of a group B worker, employers prefer recruiting group A workers to group B workers if the perceived  $r'_j \eta_i$ 's are the same,  $E[r'_j \eta_i | I_{i \in A, j}] = E[r'_j \eta_i | I_{i \in B, j}]$ . To clarify this point, we provide an explanation which builds on a framework in Lundberg (1991). As employers do not have perfect knowledge about the productivity of an individual worker, it is likely that there is a discrepancy between the skill requirement of the assigned task and the actual skill level of a worker. In practice, workers are paid by the contract wage attached to their assigned task rather than by their actual productivity. The disparity between the perceived and actual productivity imposes a

<sup>10</sup> For notational brevity, the subscript  $j$  is dropped from  $\sigma_\eta^2$ .

cost on the employer. If a worker's skill level is insufficient for an assigned task, the worker will not be perfectly performing the task. The worker will thus be overpaid, causing a loss to an employer. If a worker is over-qualified for the task, the worker will be underpaid, and is likely to change jobs in the next period. This also leads to a cost. We assume that these discrepancies due to under- and over-qualification result in costs which are characterized by a quadratic loss function of the gap between the actual productivity and wages. The expected cost for employer  $j$  of recruiting worker  $i$  under (4) is then

$$E[(P_{ij} - E[P_{ij} | I_{ij}])^2] = \left( \frac{\sigma_\xi^2}{\sigma_\eta^2 + \sigma_\xi^2} \right)^2 \sigma_\eta^2 + \left( \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\xi^2} \right)^2 \sigma_\xi^2. \quad (8)$$

As the noise of a group B worker is larger than that of a group A worker,  $\sigma_{A\xi}^2 < \sigma_{B\xi}^2$ , we can easily verify that the mismatch cost is greater when hiring a group B worker than a group A worker:

$$E[(P_{ij} - E[P_{ij} | I_{i \in A, j}])^2] < E[(P_{ij} - E[P_{ij} | I_{i \in B, j}])^2] \text{ for any } j.$$

As employers have an incentive to avoid such worker misallocation, more group A workers than group B workers are likely to be employed when they are expected to be equally productive.

Employer  $j$  always has an incentive to use group membership as a recruiting device because not using the information results in a larger cost. Not using the group membership information results in a cost of

$$E[(P_{ij} - E[P_{ij}])^2] = \sigma_\eta^2.$$

Since group A workers are preferred to group B workers who have the same observable characteristics other than group membership, we expect that the unemployment rate of group A is likely to be lower than that of group B at the time of labor market entry. The discussion in this section leads to Proposition 1.

**Proposition 1.** When employers statistically discriminate against group B workers in comparison with observationally equivalent group A workers, the group B unemployment rate will be larger than the group A unemployment rate at the time of labor market entry.



## 2.4. Experience, Employer Learning, and Statistical Discrimination

Suppose that worker  $i$  accepts an offer from employer  $j$ . Now, worker  $i$  produces an output,  $Q_{ijt}$ , at each experience level  $t=1,2,\dots,T$ . Researchers, however, do not observe these outcomes. The output,  $Q_{ijt}$ , net of the deterministic term,  $r'H_{it}$ , is a proxy for the weighted ability,  $r'_j\eta_i$ , of the worker:

$$\begin{aligned} q_{ijt} &\equiv Q_{ijt} - r'H_{it} \\ &= r'_j\eta_i + \varepsilon_{ijt}, \text{ for } t=1,2,\dots,T, \end{aligned}$$

where  $\varepsilon_{ijt}$ 's are *iid* normal random variables independent of  $\xi_{ij}$  with  $E[\varepsilon_{i \in A, jt}] = E[\varepsilon_{i \in B, jt}] = 0$  and  $Var(\varepsilon_{i \in A, jt}) = Var(\varepsilon_{i \in B, jt}) = \sigma_\varepsilon^2$ . The distribution of  $\varepsilon_{ijt}$  is common knowledge. Since we assume that learning on  $Q_{ijt}$  and  $H_{it}$  takes place publicly, the entire market learns about the realized stochastic part of output,  $q_{ijt}$ .

After observing  $q_{ij1}, q_{ij2}, \dots, q_{ijT}$ , employer  $j$  subsequently updates her initial evaluation about  $\eta_i$  of worker  $i$  in each period. The proof below is similar to that in Pinkston (2006).

$$\begin{aligned} E[P_{ijt} | I_{ij}, q_{ij1}, \dots, q_{ijT}] &= r'H_{it} + E[r'_j\eta_i | r'_j\eta_i + \xi_{ij}, r'_j\eta_i + \varepsilon_{ij1}, \dots, r'_j\eta_i + \varepsilon_{ijT}] \\ &= r'H_{it} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \frac{\sigma_\varepsilon^2 \sigma_\xi^2}{\sigma_\varepsilon^2 + T\sigma_\xi^2}} \left( \frac{\sigma_\varepsilon^2 S_{ij} + \sigma_\xi^2 \sum_{t=1}^T q_{ijt}}{\sigma_\varepsilon^2 + T\sigma_\xi^2} \right). \end{aligned} \quad (9)$$

As workers become more experienced, employer  $j$  learns more about their productivity, and the distribution of evaluations conditional on signals approaches the true productivity distribution since

$$Var(E[P_{ij} | I_{ij}, q_{ij1}, \dots, q_{ijT}]) = \frac{\sigma_\eta^4}{\sigma_\eta^2 + \frac{\sigma_\varepsilon^2 \sigma_\xi^2}{\sigma_\varepsilon^2 + T\sigma_\xi^2}} \quad (10)$$

approaches  $\sigma_\eta^2$  from below as  $T$  becomes larger. More importantly, the amount of learning is greater for group B workers than for group A workers:

$$0 < \frac{d}{dT} Var(E[P_{ij} | I_{i \in A, j}, q_{ij1}, \dots, q_{ijT}]) < \frac{d}{dT} Var(E[P_{ij} | I_{i \in B, j}, q_{ij1}, \dots, q_{ijT}])). \quad (11)$$

Intuitively, since there was less known about group B workers than group A workers initially, there is more to learn about group B workers than group A workers. Inequality (11) indicates that employers learn more about group B workers than

group A workers, implying that the mismatch cost associated with recruiting group B workers reduces faster than that associated with recruiting group A workers. Therefore, the group B unemployment rate will decrease faster than the group A unemployment rate. If workers could work forever, their productivity would be revealed perfectly and the two unemployment rates would converge to the same level. The above predictions are summarized in Proposition 2.

**Proposition 2.** When employers statistically discriminate against group B workers in comparison with observationally equivalent group A workers, and employers learn about the productivity of workers as they accumulate more experience, the unemployment rate gap between groups A and B will decrease with experience.

The logic of Proposition 2 also holds for a worker who is not continuously employed. While employed, a worker may change his job if the match quality of a new offer is better than the current match quality. The work history of the worker will be available to new employers and employer learning continues. Therefore, Proposition 2 is not affected by job-to-job movements. When a worker is unemployed, however, employer learning is interrupted. There are possibly two reasons for worker  $i$  to be unemployed at time  $t$ . First, worker  $i$  may have never been employed. The standard search model applies in this case. He still has a chance to find a job in the next period as different employers may weigh his vector of skills differently and a new signal will be drawn from (4).<sup>11</sup> Second, worker  $i$  may have been laid off by employer  $j$  and has not found a new job since then. Worker  $i$  may find another employer who offers a lower wage in the following period or who favorably weights his skills. In either of the above-mentioned cases, worker  $i$  has a chance to be employed, and once employed, employers learn according to (9).

The predictions of Propositions 1 and 2 are in line with Neumark (2012). In his paper, the difference between groups in the variance of unobservables conditional observables plays a key role in identifying the effect of group membership. Any nonzero effect may be interpreted as a result of discrimination. We exploit the prediction that the difference between groups in the variances of unobservables at the time of labor market entry conditional on observables will narrow with labor market experience in the presence of employer learning.

Using (10), one can also verify that

---

<sup>11</sup> A worker who has never been employed may have a distribution of  $\varepsilon_{ijt}$  different from that of a worker who is currently employed. This possibility is not allowed in the model, but the key theoretical implications would remain unaffected.

$$0 > \frac{d^2}{dT^2} \text{Var}(E[P_{ij} | I_{i \in A, j}, q_{ij1}, \dots, q_{ijT}]) > \frac{d^2}{dT^2} \text{Var}(E[P_{ij} | I_{i \in B, j}, q_{ij1}, \dots, q_{ijT}]) . (12)$$

While inequality (11) shows that the difference between groups in the amount of known by employers decreases with experience, inequality (12) shows that the narrowing of the difference between groups slows down with worker experience. This prediction is consistent with the empirical findings in Lange (2007). He shows that employers learn fast, meaning that most learning occurs during the early career of a worker. Inequalities (11) and (12) lead to Proposition 2'.

**Proposition 2'.** When employers statistically discriminate against group B workers in comparison with observationally equivalent group A workers, and employers learn about the productivity of workers as they accumulate more experience, the unemployment rate gap between groups A and B will decrease (at a decreasing rate) with experience.

### III. Empirical Findings

#### 3.1. Evidence of Discrimination from Descriptive Statistics

Our sample is drawn from the Integrated Public Use Microdata Series (IPUMS) March Current Population Survey (CPS) for 1971-2016.<sup>12</sup> We focus on non-Hispanic white and African-American men between the ages of 15 and 60, but exclude individuals in the non-civilian labor force and those living in group quarters. Foreign-born individuals are excluded when possible.<sup>13</sup> We obtain the measure of experience by age minus five and years of schooling.<sup>14</sup>

As a result of this derivation, about 0.4% of the sample has zero or negative experience, and these observations are excluded from the analyses.

Table 1 reports unemployment rates by race and by completed years of schooling at different experience levels during the sample period. Completed years of schooling are grouped into one of four education categories: below high school (9-11 years of schooling), high school (12 years of schooling), some college (13-15 years of schooling), and BA degree or above (16-20 years of schooling). The patterns are consistent with the predictions made by the statistical discrimination and employer

<sup>12</sup> 1971 is the first year when the information on Hispanic origin is collected by the CPS.

<sup>13</sup> Information on birth country becomes available in 1994.

<sup>14</sup> An alternative measure of experience is actual experience, although it cannot be calculated by the variables in the CPS. Altonji and Pierret (2001) employ both experience measures using the NLSY79 to examine employer learning and statistical discrimination. They find that the two sets of estimates are qualitatively the same.

learning hypotheses. First, as Proposition 1 predicts, the unemployment rate at the time of labor market entry is higher for African-Americans or the less educated than non-Hispanic whites or the more educated, respectively. There is a 12.8%p unemployment rate gap between non-Hispanic whites and African-Americans when they have 1-10 years of experience. Between high school graduates and those with a BA degree or above, the unemployment rate gap with experience up to 10 years is 9.9%p. Second, both the unemployment rate gaps by race and education decrease at a decreasing rate with experience, consistent with the predictions made by Proposition 2'.<sup>15</sup> While the descriptive statistics in Table 1 support the statistical discrimination and employer learning hypotheses, we need to verify that the patterns in Table 1 are not explained by alternative hypotheses. In addition, we need more formal analyses than simply relying on descriptive statistics to test Propositions 1 and 2'. These issues are discussed in turn in the following two subsections.

[Table 1] Unemployment Rates by Race and Education by Labor Market Experience

Experience:	1-10	11-20	21-30	31-40	Total	Observations
[B] African-American	0.225	0.113	0.088	0.086	0.138	169,288
[W] non-Hispanic White	0.097	0.049	0.040	0.041	0.059	1,435,158
Gap between [B] and [W]	-0.128	-0.064	-0.048	-0.045	-0.079	
[<HS] Below High School	0.216	0.147	0.100	0.071	0.161	257,983
[HS] High School	0.130	0.073	0.056	0.051	0.079	555,998
[SC] Some College	0.076	0.048	0.040	0.043	0.054	390,151
[BA+] BA Degree or Above	0.031	0.020	0.021	0.026	0.024	400,314
Gap between [HS] and [BA+]	-0.099	-0.053	-0.035	-0.025	-0.055	

### 3.2. Empirical Specification and Findings

As an initial step to examine statistical discrimination in the presence of employer learning, we specify an equation which can best describe the implications made by Propositions 1 and 2'. Consider a linear probability model given by

$$\begin{aligned}
 emp_{it} = & \beta_0 + \beta_r race_i + \beta_s schooling_i \\
 & + (\gamma_0 + \gamma_r race_i + \gamma_s schooling_i) \cdot experience_{it} \\
 & + (\delta_0 + \delta_r race_i + \delta_s schooling_i) \cdot experience_{it}^2 + byear_i + \varepsilon_{it}
 \end{aligned} \tag{13}$$

where  $emp_{it}$  is an indicator of employment taking on a value of 1 when employed

<sup>15</sup> The unemployment rate gaps decrease with experience, but do not converge to zero. This suggests that statistical discrimination is not the only reason for the unemployment gaps. This possibility is discussed later in the "Discussion and Alternative Explanations" section.

and 0 when unemployed,  $race_i$  is an indicator of being an African-American,  $schooling_i$  is the years of schooling,  $experience_{it}$  is the potential experience,  $byear_i$  is the birth year, and  $\varepsilon_{it}$  is an idiosyncratic error term.<sup>16</sup> We test Proposition 1 by investigating whether the so-called less-likely-to-be-employed group workers (African-Americans or less educated individuals) are less likely to be employed than the so-called more-likely-to-be-employed group workers (non-Hispanic whites or more educated individuals) at the time of labor market entry. Proposition 1 implies a negative  $\beta_r$  and a positive  $\beta_s$ . We then test Proposition 2, whether the unemployment rates of workers from the less-likely-to-be-employed group decrease at a faster rate than those of workers from the more-likely-to-be-employed group. Proposition 2 implies a positive  $\gamma_r$  and a negative  $\gamma_s$ . Finally, we examine Proposition 2', whether the speed of narrowing in the unemployment rate gap decreases with experience. Proposition 2' implies a negative  $\delta_r$  and a positive  $\delta_s$ .

The results are presented in Table 2. Column (1) tests Propositions 1 and 2 with respect to race. First, there exists an initial 13.0%p black-white unemployment rate gap. Second, the black-white unemployment gap narrows by 0.30%p for every one year increase in work experience. For example, the black-white gap in the unemployment rates is 10.0%p among workers with 10 years of experience and is 7.0%p among workers with 20 years of experience. In sum, the results suggest that there is statistical discrimination on the basis of race. Column (2) includes quadratic experience terms into the model. We find a negative estimate for  $black \times experience^2 / 100$ , while the estimates for  $black$  and  $black \times experience / 10$  do not change qualitatively. As the magnitude of the  $black \times experience / 10$  estimate is much larger than that of  $black \times experience^2 / 100$ , we can verify that the black-white unemployment is decreasing at a decreasing rate with experience. For example, the black-white gap in the unemployment rates among workers who enters the labor market for the first time is 17.1%p. Among workers with 10 years of experience, the black-white gap in the unemployment rates is 9.4%p ( $= 17.1 - 0.93 \times 10 + 0.016 \times 100$ ). Among workers with 20 years of experience, the black-white gap is 4.9%p ( $= 17.1 - 0.93 \times 20 + 0.016 \times 400$ ). Thus, the gap is decreasing at a decreasing rate. Proposition 20 is empirically supported by column (2).

Columns (3) and (4) examine statistical discrimination on the basis of education. The education variable in this analysis is obtained by subtracting 12 from the number of years of schooling and then dividing it by 4. In effect, this variable takes

<sup>16</sup> The birth year dummy variables enter the model to account for differences in employment rates between cohorts of workers. For example, a 40-year-old African-American high school graduate in 1971 likely faced very different opportunities at labor market entry from a 40-year-old worker of the same group in 2016. Calendar year dummy variables are not included because calendar year effects will be averaged out. The estimation results are not affected qualitatively and remain numerically similar after controlling for calendar year.

on a value of zero for 12 years of schooling (high school graduates) and one for 16 years of schooling (BA degree individuals). Since the years of schooling in our sample range from 9 to 20, the education variable takes on values between -0.75 and 2. In columns (3) and (4), the initial unemployment gap between high school graduates and those who have a BA degree is 10.3-11.8%p. Negative education  $\times$  experience/10 estimates are consistent with Proposition 2. For example, the results in column (3) imply that the unemployment gap between high school graduates and those who have a BA degree decreases by 0.25%p for each year of experience. A positive education  $\times$  experience<sup>2</sup>/100 estimate in column (4) indicates that the speed of learning slows down with experience, a prediction made by Proposition 20. In sum, we find evidence of statistical discrimination on the basis of education.

[Table 2] Linear Probability Model Estimates from the 1971-2016 CPS (Dependent Variable = 1 if Employed, 0 if Unemployed)

	(1)	(2)	(3)	(4)	(5)	(6)
Black	-0.130*** (0.002)	-0.171*** (0.003)			-0.116*** (0.002)	-0.160*** (0.003)
Education			0.103*** (0.001)	0.118*** (0.001)	0.099*** (0.001)	0.114*** (0.001)
Experience/10	0.015*** (0.000)	0.065*** (0.001)	0.025*** (0.000)	0.067*** (0.001)	0.022*** (0.000)	0.060*** (0.001)
Black $\times$ Experience/10	0.030*** (0.001)	0.093*** (0.004)			0.028*** (0.001)	0.095*** (0.004)
Education $\times$ Experience/10			-0.025*** (0.000)	-0.053*** (0.001)	-0.025*** (0.000)	-0.053*** (0.001)
Experience <sup>2</sup> /100		-0.013*** (0.000)		-0.011*** (0.000)		-0.010*** (0.000)
Black $\times$ Experience <sup>2</sup> /100		-0.016*** (0.001)				-0.017*** (0.001)
Education $\times$ Experience <sup>2</sup> /100				0.008*** (0.000)		0.008*** (0.000)
Birth Year	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,370,855	1,370,855	1,370,855	1,370,855	1,370,855	1,370,855
R-squared	0.024	0.029	0.037	0.039	0.044	0.046

Notes: Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

We find that each coefficient estimate in columns (5) and (6) is numerically similar to the corresponding coefficient estimate in columns (1)-(4). Overall, the results in columns (1)-(4) as well as those in (5) and (6) of Table 2 are consistent with Propositions 1 and 2'. It happens because the correlation between schooling and race is not substantial. Econometrically, not controlling for schooling in columns (1) and (2) or not controlling for race in columns (3) and (4) does not

cause a serious omitted variables bias problem. We conclude that African-American workers and low-educated workers have initially worse labor market opportunities in terms of employment probability, but that their unemployment rates improve faster with experience at a decreasing rate.

### 3.3. Discussion and Alternative Explanations

This subsection discusses whether the predictions made by the theoretical model presented in this paper can also be explained by other hypotheses such as human capital theory, taste-based discrimination, or search and matching models. We find that these alternative explanations, at least as currently developed, are not adequate for explaining the estimates presented in this paper.

With regard to human capital theory, so far in the discussion in this paper, the rate of human capital attainment has been set the same for both groups of workers. A more realistic assumption, however, is that group B workers have fewer opportunities for human capital investment than do group A workers.<sup>17</sup> Employers may offer fewer training or promotion opportunities to group B workers than to group A workers. Moreover, if African-Americans obtain more education than whites of similar cognitive ability, as Lang and Manove (2011) find, African-Americans will have fewer chances for training conditional on education. Therefore, the parameter vector  $r$  in (1) in reality is likely to be greater for group A workers than group B workers. Now suppose that there exists statistical discrimination in the opportunities for human capital investment. Enforcing  $r$  to be common across groups makes Proposition 2 or 2' more difficult to hold empirically than allowing  $r$  to be group-specific. Consequently, finding the pattern predicted by Proposition 2 or 2' in the data with the common  $r$  restriction serves as stronger evidence of statistical discrimination.

Next, we consider the theory of taste-based discrimination. This theory can explain why less-likely-to-be-employed groups have higher unemployment rates. In Table 1, there is on average a 7.9%p difference in the unemployment rates between non-Hispanic whites and African Americans and a 5.5%p difference in the unemployment rates between high school graduates and those who have a BA degree or above. The theory of taste-based discrimination, however, cannot explain why the gaps in unemployment rates between the less-likely-to-be-employed and more-likely-to-be-employed groups narrow with experience. The black-white unemployment rate gap is 12.8%p for workers with 1-10 years of experience and

---

<sup>17</sup> It is possible that disadvantaged workers receive more on-the-job training once hired. Holzer and Neumark (2000) find that African-American male workers spend more time with supervisors or co-workers in the presence of affirmative action. However, there is no clear evidence that this informal training leads to a higher rate of human capital investment. They find, however, that women receive significantly more formal on-the-job training when affirmative action is present.

4.5%p for workers with 31-40 years of experience. The unemployment rate gap between high school graduates and those who have a BA degree or above is 9.9%p for young workers and 2.5%p for the most experienced workers. In this sense, the taste-based discrimination story is insufficient to explain Proposition 2 or 2'.

**[Table 3]** Unemployment Rates by Race and Education at Different Experience Levels

Experience:	1-10	11-20	21-30	31-40	Total	Observations
<b>White</b>						
[<HS] Below High School	0.193	0.136	0.092	0.065	0.147	216,474
[HS] High School	0.115	0.066	0.051	0.047	0.071	489,335
[SC] Some College	0.068	0.044	0.037	0.040	0.048	350,320
[BA+] BA Degree or Above	0.030	0.018	0.020	0.025	0.023	379,029
Total	0.097	0.049	0.040	0.041	0.059	1,435,158
<b>Black</b>						
[<HS] Below High School	0.393	0.205	0.143	0.104	0.253	41,509
[HS] High School	0.235	0.129	0.101	0.092	0.148	66,663
[SC] Some College	0.148	0.093	0.073	0.081	0.103	39,831
[BA+] BA Degree or Above	0.061	0.043	0.037	0.050	0.047	21,285
Total	0.225	0.113	0.088	0.086	0.138	169,288
<b>Black-White Unemployment Gap</b>						
[<HS] Below High School	-0.200	-0.069	-0.051	-0.039	-0.106	
[HS] High School	-0.120	-0.063	-0.050	-0.045	-0.077	
[SC] Some College	-0.080	-0.049	-0.036	-0.041	-0.055	
[BA+] BA Degree or Above	-0.031	-0.025	-0.017	-0.025	-0.024	
Total	-0.128	-0.064	-0.048	-0.045	-0.079	

Finally, we discuss whether a search model can produce results that are consistent with the empirical results of this paper. A standard search model predicts that unemployment decreases with experience, which seems to be consistent with the patterns in Table 1. The model, however, fails to explain the patterns in Table 3, which shows that the male black-white employment gap is much larger among those less-educated than among those more-educated. For each of the four education groups, the black-white difference decreases with experience. This tendency is larger for the less educated. For those who have 9-11 years of schooling, the black-white unemployment rate gap is 20.0%p during the first 10 years of work experience. This gap diminishes with experience, reaching 3.9%p with 31-40 years of labor market experience. For the highly educated, the gap stops diminishing relatively early in a career. The initial unemployment gap is 3.5%p, decreasing to 1.7%p by 21-30 years of experience, and rising to 2.5%p afterwards. In general, initially the black-white unemployment gap decreases with education, but the improvement in the black-white unemployment gap is slower for the highly educated. In the absence of statistical discrimination, we would expect that the variance of perceived productivity conditional on the signals in (10) would not



depend on group identity. This implies that the unemployment rates by education/race group would be the same at any given experience level. This prediction is not borne out in Table 3.

A search model accompanied by taste-based discrimination may explain these patterns, but possibly after making many restrictions. Lang and Lehmann (2012) also notice this pattern. They state that this pattern is not explained by an existing search model, even though there has been an increase in research in this field. They point out that while taste-based discrimination models can generate wage and unemployment duration differentials between group A workers and group B workers when combined with search, no existing model for taste-based discrimination can explain the unemployment differential. Whether a search model can generate the patterns found in this paper is still an open question.

## IV. Robustness

### 4.1. Trends in Statistical Discrimination over Time

Table 2 covers the CPS sample from 1971-2016, but analyzing the whole sample period as one may hide variations of underlying feature over time. There are two reasons for this. The first is a possible time trend in statistical discrimination. Over these years, the degree of statistical discrimination may have shifted if there were any changes in employers' information sets. For example, the diversification of college majors or the reduced cost of information acquisition due to internet access may provide richer information to potential employers. In case of improvements in screening ability, the degree of statistical discrimination may phase out when (13) is estimated by calendar year. Second, as the information on birth place becomes available from the 1994 CPS, the sample used to estimate the results in Table 2 excludes foreign-born individuals from 1994. It is necessary to check whether the different sample selection criteria used from 1994 has led to biased results.

To see whether statistical discrimination has changed over time, we split our sample period into four subperiods: 1971-1982, 1982-1993, 1994-2005, and 2005-2016. The 1982 and 2005 samples are included in two of the sub-periods so that all four sub-periods are of equal length. Overall, we do not find any noticeable changes in the statistical discrimination and employer learning processes over time. As we see in Table 4, the coefficients of  $\text{black} \times \text{experience}/10$ ,  $\text{education} \times \text{experience}/10$ ,  $\text{black} \times \text{experience}^2/100$ , and  $\text{education} \times \text{experience}^2/100$  remain stable over the four subperiods. We interpret the results in columns (1), (3), (5), and (7). The initial black-white unemployment rate gap is 10.7-11.4%p and the gap narrows by 0.22%p per year of experience. The initial unemployment rate gap between high

[Table 4] Estimates from Subsamples of the 1971-2016 CPS (Dependent Variable = 1 if Employed, 0 if Unemployed)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1971-1982	1971-1982	1982-1993	1982-1993	1994-2005	1994-2005	2005-2016	2005-2016
Black	-0.107*** (0.002)	-0.156*** (0.004)	-0.114*** (0.002)	-0.159*** (0.004)	-0.109*** (0.003)	-0.149*** (0.005)	-0.110*** (0.004)	-0.144*** (0.006)
Education	0.095*** (0.001)	0.109*** (0.001)	0.102*** (0.001)	0.118*** (0.002)	0.093*** (0.001)	0.116*** (0.002)	0.100*** (0.002)	0.130*** (0.003)
Experience/10	0.022*** (0.000)	0.062*** (0.001)	0.024*** (0.000)	0.057*** (0.001)	0.008*** (0.001)	0.039*** (0.002)	0.010*** (0.001)	0.046*** (0.004)
Black × Experience/10	0.025*** (0.001)	0.092*** (0.004)	0.026*** (0.001)	0.090*** (0.004)	0.024*** (0.001)	0.082*** (0.005)	0.021*** (0.002)	0.071*** (0.007)
Education × Experience/10	-0.023*** (0.000)	-0.050*** (0.001)	-0.024*** (0.000)	-0.052*** (0.002)	-0.023*** (0.001)	-0.058*** (0.002)	-0.023*** (0.001)	-0.060*** (0.003)
Experience <sup>2</sup> /100		-0.010*** (0.000)		-0.009*** (0.000)		-0.008*** (0.000)		-0.009*** (0.001)
Black × Experience <sup>2</sup> /100		-0.017*** (0.001)		-0.016*** (0.001)		-0.015*** (0.001)		-0.012*** (0.002)
Education × Experience <sup>2</sup> /100		0.007*** (0.000)		0.007*** (0.000)		0.009*** (0.000)		0.009*** (0.001)
Birth Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,281,592	1,281,592	1,062,978	1,062,978	704,048	704,048	381,291	381,291
R-squared	0.036	0.039	0.041	0.043	0.044	0.046	0.048	0.049

Notes: Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

school graduates and university graduates is in the range of 9.3-10.2%p without any specific trend. The gap decreases by 0.23%p per year of experience. In sum, there is no substantial time effect on the employer learning process. Also, the effect of including and excluding foreign-born workers is not critical.<sup>18</sup>

## 4.2. Heterogeneity in Statistical Discrimination by Education

From Table 3, we have found that statistical discrimination is stronger for less educated individuals and weaker for more educated individuals. We investigate this issue further. In general, more educated individuals are more specialized, and specialization makes it easier for employers to observe worker productivity. Due to (8), the mismatch cost is smaller among highly educated individuals. This implies that the degree of statistical discrimination on the basis of race is smaller for more educated workers than less educated workers. In addition, since much is known about the productivity of highly educated workers at the beginning of their careers, the amount of learning is smaller and their productivity can be ascertained in a shorter period than for less educated workers. To verify this conjecture, we examine Propositions 1 and 2' by completed years of schooling. Table 5 reports these results. We find that the initial unemployment gap is smaller for the more educated and that learning is faster for the low educated. In column (3), among high school graduates, the initial racial unemployment gap is 12.0%p and is decreasing by 0.27%p per year of experience. In column (7), among individuals with a BA degree or higher, the initial racial unemployment gap is much lower, 3.2%p, and the gap decreases at a much slower rate of 0.04%p per year of experience. In other words, the productivity of African-American university graduates is almost as predictable as that of white university graduates.

## 4.3. Evidence from the NLSY79

We examine Propositions 1 and 2' using a sample drawn from the NLSY79 for 1979-2010. This exercise is useful in several respects. First of all, previous research has found evidence for statistical discrimination using the NLSY79. We can verify whether the empirical findings are robust to other data sets and compare our findings on employment with those of previous papers on wage discrimination. Second, the NLSY79 individuals were born within a narrow period of 1957-1964. This alleviates concerns about cohort effects due to cohort heterogeneity.

---

<sup>18</sup> We know from other sources that the share of foreign-born individuals did not exceed 15% during the pre-1994 period.

[Table 5] Linear Probability Model Estimates from the CPS by Completed Years of Schooling (Dependent Variable = 1 if Employed, 0 if Unemployed)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<HS	<HS	HS	HS	SC	SC	BA+	BA+
Black	-0.213*** (0.006)	-0.256*** (0.008)	-0.120*** (0.004)	-0.156*** (0.006)	-0.081*** (0.004)	-0.107*** (0.006)	-0.032*** (0.004)	-0.040*** (0.006)
Experience/10	0.019*** (0.001)	0.017*** (0.003)	0.016*** (0.000)	0.059*** (0.002)	0.007*** (0.000)	0.032*** (0.002)	0.001*** (0.000)	0.025*** (0.001)
Black × Experience/10	0.060*** (0.002)	0.148*** (0.009)	0.027*** (0.001)	0.081*** (0.006)	0.016*** (0.002)	0.054*** (0.007)	0.004*** (0.002)	0.016*** (0.007)
Experience <sup>2</sup> /100		0.001 (0.001)		-0.011*** (0.000)		-0.006*** (0.000)		-0.006*** (0.000)
Black × Experience <sup>2</sup> /100		-0.024*** (0.002)		-0.014*** (0.001)		-0.010*** (0.002)		-0.003* (0.002)
Birth Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	152,392	152,392	498,332	498,332	338,960	338,960	381,171	381,171
R-squared	0.046	0.047	0.024	0.027	0.010	0.012	0.002	0.004

Notes: Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

The NLSY79 sample used in this analysis consists of the cross-sectional sample of whites and blacks and the oversample of blacks. Under this approach, an individual is classified as unemployed in a given year if he was in the labor force for at least 26 weeks of that year and was unemployed for at least half of that period while in the labor force. This measure is constructed using three variables: the number of weeks employed, unemployed, and out of the labor force. Observations lacking any of the three variables are dropped. In addition, of the remaining observations, the three variables do not add up to 52 in about 3% of cases and these cases are also excluded from the analyses.

**[Table 6]** Linear Probability Model Estimates from the NLSY79 (Dependent Variable = 1 if Employed, 0 if Unemployed)

	(1)	(2)	(3)	(4)	(5)	(6)
Black	-0.112*** (0.005)	-0.164*** (0.010)			-0.098*** (0.005)	-0.148*** (0.009)
Education			0.089*** (0.004)	0.114*** (0.006)	0.076*** (0.004)	0.097*** (0.006)
Experience/10	0.011*** (0.001)	0.058*** (0.005)	0.023*** (0.001)	0.099*** (0.006)	0.018*** (0.001)	0.073*** (0.006)
Black × Experience/10	0.021*** (0.003)	0.102*** (0.012)			0.018*** (0.003)	0.097*** (0.012)
Education × Experience/10			-0.018*** (0.002)	-0.053*** (0.009)	-0.016*** (0.002)	-0.044*** (0.008)
Experience <sup>2</sup> /100		-0.015*** (0.001)		-0.023*** (0.002)		-0.016*** (0.002)
Black × Experience <sup>2</sup> /100		-0.023*** (0.003)				-0.023*** (0.003)
Education × Experience <sup>2</sup> /100				0.009*** (0.003)		0.007*** (0.003)
Birth Year	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	62,822	62,822	62,822	62,822	62,822	62,822
R-squared	0.029	0.035	0.027	0.031	0.046	0.052

Notes: Robust standard errors in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6 presents the estimation results of (13) using the NLSY79, where the dependent variable is the indicator of employment. Overall, the NLSY79 estimates are qualitatively the same as the CPS estimates in Table 2. In column (1), an initial 11.2%p black-white gap in the unemployment rates narrows down by 0.21%p per year of experience. In column (2), where the quadratic terms are included, there is an initial 16.4%p unemployment rate gap which decreases at a decreasing rate. In column (3), the initial unemployment rate gap between high school graduates and those who have a BA degree is 8.9%p, and the gap decreases by 0.18%p per year of

experience. When the quadratic terms are included in column (4), we find that the initial unemployment rate gap decreases at a decreasing rate with experience. Overall, the signs of the estimates in columns (1)-(4) as well as those in columns (5)-(6) are consistent with Propositions 1 and 2'.

## V. Concluding Remark

This paper proposes a novel approach to identify statistical discrimination using variables commonly available in usual data sets. The proposed strategy exploits the heterogeneity in employers' acquisition processes for information on worker productivity. While the process is worker-employer specific in general, employers can evaluate specific groups of workers systematically better than the other groups of workers. This information structure enables us to show that even when an employer evaluates the productivity of two applicants as equal, the employer prefers recruiting the applicant who belongs to a group which can be evaluated with more precision. Over time employers observe workers' performance repeatedly, and this learning process will depend on a worker's group membership. If these predictions hold, workers in a group with noisier signals are more likely to be unemployed than workers in a group with cleaner signals at any given experience level and observable characteristics, but as workers gain experience the noisier signal group's unemployment rate will decrease faster than that of the cleaner signal group.

The empirical findings using the March CPS are consistent with the theoretical predictions. Initially, black workers are more likely to be unemployed than white workers when they are young, but the black-white gap in unemployment rates narrows with experience at a given educational level. Similarly, education is negatively correlated with the probability of becoming unemployed, but the unemployment rates of low-educated workers drop faster than those of highly educated workers with experience for any given race. In sum, this paper provides a theoretical foundation for one of the well-known empirical regularities that the black-white employment gap is larger among low-skilled workers than among high-skilled ones.

## References

- Altonji, J. G. (2005), “Employer Learning, Statistical Discrimination and Occupational Attainment,” *American Economic Review (Papers and Proceedings)*, 95(2), 112–117.
- Altonji, J. G. and C. R. Pierret (2001), “Employer Learning and Statistical Discrimination,” *Quarterly Journal of Economics*, 116(1), 313–350.
- Cho, J., H. Chu, H. E. Kim, and J. Lee (2016), “Productivity Distribution and Economic Growth,” *Korean Economic Review*, 32(1), 23–40.
- Cornell, B. and I. Welch (1996), “Culture, Information, and Screening Discrimination,” *Journal of Political Economy*, 104(3), 542–571.
- Flood, S., M. King, S. Ruggles, and J. R. Warren (2015), *Integrated Public Use Microdata Series, Current Population Survey: Version 4.0*. [dataset]. Minneapolis: University of Minnesota. <http://doi.org/10.18128/D030.V4.0>.
- Holzer, H. J. and D. Neumark (2000), “What Does Affirmative Action Do?” *Industrial and Labor Relations Review*, 53(2), 240–271.
- Lang, K. and M. Manove (2011), “Education and Labor Market Discrimination,” *American Economic Review*, 101(3), 1467–1496.
- Lang, K. and J. K. Lehmann (2012), “Racial Discrimination in the Labor Market: Theory and Empirics,” *Journal of Economic Literature*, 50(4), 959–1006.
- Lange, F. (2007), “The Speed of Employer Learning,” *Journal of Labor Economics*, 25(1), 651–691.
- Lundberg, S. J. and R. Startz (1983), “Private Discrimination and Social Intervention in Competitive Labor Markets,” *American Economic Review*, 73(3), 340–347.
- Moro, A. (2003), “The Effect of Statistical Discrimination on Black-White Wage Inequality: Estimating a Model with Multiple Equilibria,” *International Economic Review*, 44(2), 467–500.
- Neal, D. (2004), “The Measured Black-White Wage Gap among Women Is Too Small,” *Journal of Political Economy*, 112(1), S1–28.
- Neumark, D. (1999), “Wage Differentials by Race and Sex: the Roles of Taste Discrimination and Labor Market Information,” *Industrial Relations*, 38(3), 414–442.
- \_\_\_\_\_ (2012), “Detecting Discrimination in Audit and Correspondence Studies,” *Journal of Human Resources*, 47(4), 1128–1157.
- Oettinger, G. S. (1996), “Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap,” *Journal of Labor Economics*, 14(1), 52–78.
- Phelps, E. S. (1972), “The Statistical Theory of Racism and Sexism,” *American Economic Review*, 62(4), 659–661.
- Pinkston, J. C. (2003), “Screening Discrimination and the Determinants of Wages,” *Labour Economics*, 10(6), 643–658.
- \_\_\_\_\_ (2006), “A Test of Screening Discrimination with Employer Learning,” *Industrial and Labor Relations Review*, 59(2), 267–284.
- Ritter, J. A. and L. J. Taylor (2011), “Racial Disparity in Unemployment,” *Review of Economics and Statistics*, 93(1), 30–42.