The Boss is Watching: How Monitoring Decisions Hurt Black Workers*

Costas Cavounidis† Kevin Lang‡ Russell Weinstein§

August 6, 2020

Abstract

African Americans face shorter employment durations than similar whites. We hypothesize that employers discriminate in acquiring or acting on ability-relevant information. In our model, monitoring black but not white workers is self-sustaining: new black hires are more likely to have been screened by a previous employer, causing firms to discriminate in monitoring. We confirm our predictions that the layoff hazard is initially higher for black workers but converges to that for white workers, and declines with AFQT more sharply for black workers. Two additional predictions, lower lifetime incomes and longer unemployment durations for black workers, are known to be empirically supported.

*This research was funded in part by NSF grants SES-1260917 and SES-1851636. It draws on similar intuition to Cavounidis and Lang (2015) but greatly simplifies the theory and adds an empirical section testing the main new predictions. We are grateful to Cheonghum Park for excellent research assistance and to Gautam Bose, Sambuddha Ghosh, Larry Katz, Ed Lazear, Calvin Luscombe, Ioana Marinescu, Derek Neal, Marina Soley-Bori, Bill Spriggs and participants in seminars and conferences at Boston University, Cornell, CUHK-Shenzhen, EALE, IZA/OECD, NBER, Tufts, the University of California, Riverside, SOLE, the University of Oklahoma, UQAM and West Point Military Academy for helpful comments and suggestions on this and the earlier paper. The usual caveat applies.

†University of Warwick, c.cavounidis@warwick.ac.uk

‡Boston University, NBER and IZA, lang@bu.edu

§University of Illinois at Urbana-Champaign and IZA, weinst@illinois.edu

[T]he Black screwup ... faces the abyss after one error, while the White screwup is handed second chances .... - Ibram X. Kendi, *How to Be an Antiracist*, New York, NY: One World, 2019, p.93.

1 Introduction

Many Americans, especially African Americans, believe black workers 'don't get second chances'[1] or that they face additional scrutiny in the workplace. Similarly, black workers are admonished to be 'twice as good'[2] in order to succeed. If black workers are subject to higher standards or scrutinized more heavily, we expect this to be reflected in more separations.

Indeed, the data support the idea of shorter employment duration for black workers.[3] Bowlus, Kiefer and Neumann (2001) detect and ponder the disparity in job destruction rates; Bowlus and Eckstein (2002) estimate that young black male high school graduates had roughly 2/3 the job spell duration of their white counterparts.[4] In addition, more of their job spells end in unemployment, suggesting that black workers have much shorter employment spells. Both papers assume an exogenously higher separation rate for black workers to fit their models to the data. Lang and Lehmann (2012) show that differences in unemployment duration alone are insufficient to account for the black/white unemployment rate gap and therefore that black workers' employment stints are shorter. This aspect of labor discrimination has thus far eluded theoretical explication.

In this paper, our proposed explanation for differential employment durations is, in its broadest sense and consistent with the aforementioned observations, that firms discriminate in the acquisition or use of productivity-relevant information. That is, firms either learn differently about black workers or, when information regarding ability is received, they condition how they act on it on workers' race. Crucially, we establish that such discrimination can be self-perpetuating.

The essence of our model is that, because black workers are more closely scrutinized, a larger share of low-performance workers will separate into unemployment. As a result, since productivity is correlated across jobs, the black unemployment pool is 'churned' and therefore weaker than the white unemployment pool. Since workers can, at least to some extent, hide

[1] This assertion can be found in a range of occupations including football coaching (Reid, 2015), music and films (*The Guardian*, 2014) as well as more generally (Spencer, 2014).

[2] Coates (2012) and Mabry (2012)

[3] Throughout this paper we distinguish between employment duration by which we mean the length of an employment spell and job duration by which we mean the time a worker spends with a particular employer. Job duration depends on, among other factors, the arrival rate of outside offers. Our model abstracts from job-to-job transitions, but can incorporate them without trouble, as shown in Section 4.6.8.

[4] Using the NLSY data for 1985 and 1988.

their employment histories, race serves as an indicator of expected worker productivity. This in turn makes monitoring newly hired black (but not white) workers optimal for firms. Figure 1 illustrates employment in the two labor markets. The churning mechanism is shared with Masters (2014), where information acquisition takes the form of exogenous pre-employment signals rather than endogenous monitoring on the job.

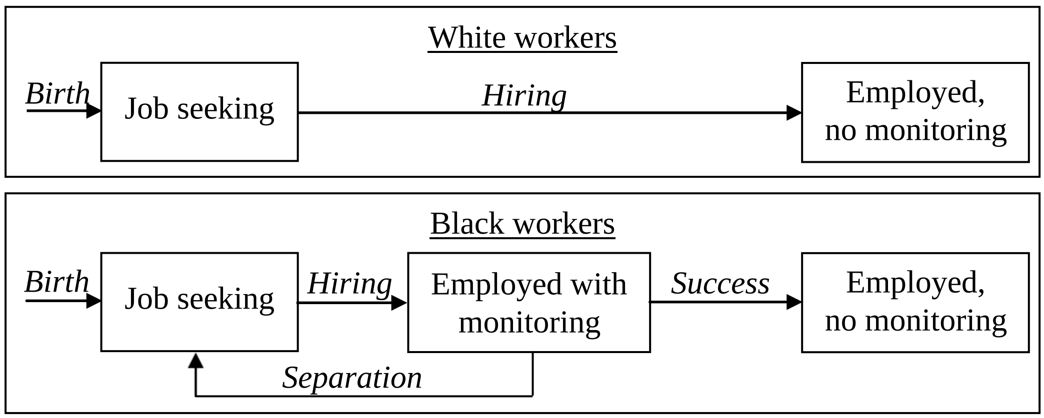


Figure 1: White workers' perpetual employment, black workers' churning cycle

Our empirical analysis begins by providing suggestive evidence that black workers are more heavily supervised even in similar occupations, at least if they have no more than a high school education, a condition that applies to the vast majority of black workers during the period for which we have data on supervision. We also show that higher scores on the Armed Forces Qualifying Test (AFQT), a measure of cognitive skill, reduce black unemployment more than white unemployment. This is consistent with monitoring of black but not white workers, and lower-skill black workers returning to the pool of unemployed workers after failing monitoring.

Importantly for 'testability,' our model has excess empirical content. First, it predicts that involuntary separations from employment will initially be higher for black workers than for white workers but that these hazards will converge with seniority. As seniority increases it is more likely that workers have passed monitoring, and are good matches with the firm. We test and largely confirm this previously untested prediction using the National Longitudinal Survey of Youth 1979 (NLSY79). The finding is robust to a variety of sample selection decisions, approaches to smoothing the separation hazard functions, measures of seniority and proxies for involuntary separation, and strengthens with the inclusion of controls.

In addition, our model implies that high unobserved ability will have a larger effect on reducing unemployment and layoffs among black than among white workers. Following a tradition dating from Farber and Gibbons (1996), we treat performance on the AFQT, after controlling for observables, as unobserved by the market. We show that AFQT does have a stronger negative effect on both unemployment and layoffs for black than for white workers.

There are multiple equilibria in our model, a property it shares with models of rational stereotyping or self-confirming expectations (Coate and Loury, 1993). However, in our model discrimination is not simply a product of coordination failure; instead, history matters. A group that begins with a low level of skills for which only the bad (monitoring) equilibrium exists will remain in that equilibrium even if its skill level rises to a level consistent with the existence of both the good and bad equilibria. Even if black workers are, on average, more skilled than white workers, whites can be in the good steady-state and black workers in the bad steady-state because of a history of lower access to schooling and other human capital investments. Equalizing the human capital that black and white workers bring to the labor market may be insufficient to equalize labor market outcomes. In contrast, in self-confirming expectations models, if we could just convince black workers to invest in themselves and employers that they have invested, we would immediately jump to the good equilibrium.

There is an abundance of evidence that black workers face lower wages and longer unemployment duration than white workers. Moreover, these disparities are less prevalent (and perhaps in some cases nonexistent) for the most skilled workers as measured by education or performance on the AFQT. While there are a plethora of models intended to explain wage or unemployment differentials, none addresses both and their relation to skill.[5] Since in our model newly hired black workers are on average less productive than white workers, their wages are lower and firms that expect to hire black workers anticipate less profit from a vacancy and therefore offer fewer jobs. Consequently, black workers have longer unemployment durations.

We derive additional implications from informal extensions to the model. The higher level of scrutiny increases the return to skill for black workers, consistent with evidence that black workers invest more in schooling compared with apparently equivalent white workers. In addition, if unemployment history is partially observable, black job seekers who have experienced enough turnover may be permanently relegated to low-skill, low-wage jobs. Although we do not wish to overstate the predictive power of the model, we note that until around 1940, black and white workers had similar unemployment rates (Fairlie and

[5] Many models (e.g. Aigner and Cain, 1977; Becker, 1971; Bjerk, 2008; Charles and Guryan, 2011; Coate and Loury, 1993; Fryer, 2007; Lang, 1986; Lang and Manove, 2011; Lundberg and Startz, 1983; Moro and Norman, 2004) assume market clearing and therefore cannot address unemployment patterns. Search models (e.g. Black, 1995; Bowlus and Eckstein, 2002; Lang and Manove, 2003; Lang, Manove and Dickens, 2005; Rosen, 1997) can explain unemployment differentials, but assume otherwise homogeneous workers and thus cannot address wage differentials at different skill levels. Peski and Szentes (2013) treat wages as exogenous. In general, discrimination models have not addressed employment or job duration. See the review in Lang and Lehmann (2012).

Sundstrom, 1999), while black workers faced lower wages. This is consistent with a setting in which, due to low human capital investments, black workers were assumed to have low productivity at most jobs and therefore not monitored for quality. 'Churning' of the black labor market would not begin until human capital investments were sufficiently high.

We believe that the broad implications of our model can be derived through a variety of formalizations. The key elements common to these are:

i. that a worker's productivity at different firms is correlated,

ii. that workers cannot or do not signal their ability and that they can, at least imperfectly, hide their employment histories,[6]

iii. that firms must therefore, to some degree, statistically infer worker ability,

iv. that further information about match productivity arrives during production, and is either costly, imperfect, or both, and

v. that this information, if obtained, may affect retention, so that firm behavior affects the average unemployed worker's ability.

The details of our formal model are driven by our desire for a theoretically rigorous model of wage-setting in a dynamic framework with asymmetric information. Firms and workers bargain over wages and use a costly monitoring technology to assess the quality of the match, which is correlated with the worker's underlying type.

Therefore, use of the monitoring technology depends on the firm's prior: if the belief that a worker is well-matched is sufficiently high or sufficiently low, it will not be worth investing resources to determine match quality. However, if the cost of determining the match quality is not too high, there will be an intermediate range at which this investment is worthwhile. Firm beliefs about black, but not white, workers fall in this region. Consequently, they are subject to heightened scrutiny and are more likely to be found to be a poor match and fired. The increased scrutiny ensures that the pool of unemployed black workers has a higher proportion of workers who have been found to be a poor match at one or more prior jobs. And therefore employers' expectations that black workers are more likely to be poor matches is correct in equilibrium. This, nested in a search model, generates the empirical predictions discussed above.[7]

[6] In particular, they must sometimes be able to omit or mischaracterize prior bad matches.

[7] Note that our model abstracts from moral hazard and that performance is observed objectively. MacLeod (2003) develops an interesting model in which biased subjective assessments interact with moral hazard concerns.

This churning equilibrium is hard to escape. This is disheartening since *policy succeeding at convergence of group characteristics may fail to equate labor market outcomes.* Only if the skill level of black workers is raised sufficiently above that of white workers (technically the proportion of good workers is sufficiently high), does the bad equilibrium cease to exist and white and black workers receive similar treatment.

2 The Model

2.1 Setup

There are two worker groups, 'black' and 'white'. Race is observable by the worker and employers but does not have any direct impact on production.

At all times a steady flow of new workers is born into each population group.[8] A proportion $g \in (0,1)$ of new workers are type α, for whom every job is a good match.[9] The rest, referred to as type β, have probability $\beta \in (0,1)$ of being a good match at any particular job. The probability of a worker being good at a job, conditional on her type, is independent across jobs. Worker type is private to the worker. Workers begin their lives unemployed. The probability a new worker is good at a particular job is

$$\theta_0 = g + (1-g)\,\beta. \tag{1}$$

Employers cannot directly observe worker type or employment history,[10] but can instead draw statistical inferences from race.

2.2 Match Quality

Production, the payment of wages and the use of the monitoring technology occur in continuous time using a common discount rate r.

Workers can be either well-suited to a task (a 'good' match), producing q per unit time; or ill-suited (a 'bad' match), producing expected output $q - \lambda c$ per unit time. We can interpret the lower productivity of bad workers as errors or missed opportunities, each costing the firm c, that arrive at a constant rate λ. Under this interpretation, opportunities for error are

[8] We do not allow for death but could do so at the cost of a little added complexity.

[9] Having type α workers perform well at every job is not essential to the argument but simplifies presentation significantly.

[10] At a more informal level, we believe that workers have some ability to hide their employment history and that they will not report information speaking to their own low ability. We show the model is robust to imperfect history revelation in Section 4.6.6.

also opportunities to learn the quality of the match as well-matched workers are observed to avoid errors.[11]

The employer does not know the match quality without monitoring. During production, the firm may use a technology that may produce a fully informative signal about match quality. If the signal shows the match to be bad, the firm may terminate it immediately, receiving 0.

In keeping with the opportunities-for-errors interpretation, we assume the signal arrives at a constant hazard rate λ. The monitoring technology costs b per unit time, so that the expected cost of information is $\int_0^\infty be^{-\lambda t}dt = b/\lambda$ and its expected discounted cost is $\int_0^\infty (e^{-rt}b)e^{-\lambda t}dt = b/(\lambda + r)$. The principal benefit of a signal whose arrival is exponentially distributed, rather than one that arrives deterministically, is that it makes the employment survival function more realistic. In addition, it allows for a certain stationarity in the model: so long as no signal has arrived, the underlying incentives do not change.

For monitoring to ever be useful, matches revealed to be bad must separate. A sufficient condition for this is that $q - \lambda c < 0$. Additionally, we intend that β workers will not be willing to reveal their type in bargaining. To this end, we make the sufficient and simple assumption that such a match is unproductive, regardless of the monitoring choice:

$$\text{(C1)} \qquad \max\left\{q - (1-\beta)\lambda c,\ \beta\frac{q}{r} + (1-\beta)\frac{(q-\lambda c)}{\lambda + r} - \frac{b}{\lambda + r}\right\} \leq 0.$$

It is *much* stronger than necessary. In general, it is sufficient that any wage at which a firm would knowingly hire a β worker is low enough that the worker would rather reject it in order to rematch at a higher (pooling) wage. Assumption (C1) ensures that such separation in search of a new match is beneficial regardless of the expected duration of unemployment.

2.3 Job Search

When a worker is born or her match is terminated, she becomes unemployed. Unemployed workers are stochastically matched to firms, which occurs at a constant hazard μ. For the moment, we treat this rate as exogenous; it will be endogenized in Section 4.5 to address unemployment duration. When a match dissolves, transfers cease and the worker becomes unemployed. A firm does not recoup a vacancy and therefore receives a payoff of 0 on termination.[12]

In the unemployed state, workers merely search for new jobs; we normalize the flow util-

[11] Alternatively, we could assume that the flows are $q - d$ and q with $d \equiv \lambda c$ and that λ is the arrival rate of opportunities to measure the flows.

[12] This occurs naturally due to free entry when vacancy creation is endogenized; see Section 4.5.

ity from this state to 0. The value from unemployment is thus simply the appropriately discounted expected utility from job-finding and is invariant to history. The expected discount on job-finding is $\int_0^\infty e^{-rt}\mu e^{-\mu t}dt = \mu/(\mu + r)$; the value of a new job will depend on the equilibrium. We denote the value of the job-finding state in a market where the average worker quality is θ as U_θ^α for type α workers and U_θ^β for type β workers. These will be constant in steady-state. Furthermore, in order to have simple dynamics outside steady-state, we assume that workers are *myopic*: they bargain as though the future wage distribution is identical to the current one.

2.4 Wage Setting

Given the asymmetry in information between the worker, who knows her type, and the firm, the Nash bargaining model is unusable and the Rubinstein (1982) one suffers from a multiplicity of equilibria. If a β worker does not want to reveal her type, as follows from our assumptions, then the β worker will have to bargain as if she were an α worker. Since in this case the firm cannot distinguish with which type it is bargaining, it should act as if it were bargaining with a random draw from the unemployment pool. Thus an intuitively appealing solution is the outcome that would be reached in Nash bargaining between a firm calculating its rents on the assumption of a random draw from the pool and a worker calculating her rents as if she were an α worker.

We posit a simple wage bargaining model akin to Lauermann and Wolinsky (2016) that produces this outcome albeit only in expectation. When a worker and firm meet, a wage offer w is randomly drawn from some distribution F. They then simultaneously choose whether to accept or reject the offer. If either rejects the offer, the match is dissolved; the firm receives 0 while the worker searches for the next firm. If both parties accept the offer, production proceeds at that wage. We are looking for Perfect Bayesian Equilibria in which neither party uses a (weakly) dominated strategy.[13] Using a randomly drawn take-it-or-leave-it offer allows us to escape both the multiplicity of equilibria caused by off-path beliefs if the worker can make offers, and the Diamond paradox if only the firm makes offers.[14]

To ensure that the wage process does not end in disagreement in equilibrium, we assume only agreeable wages are proposed. Specifically, we assume that F is a uniform distribution on the set of wages the firm and worker would both accept; thus, wage negotiation is as though an arbitrator proposes any wage on the contract curve with equal likelihood. Cru-

[13] This is needed to rule out equilibria where mutually acceptable wages are rejected by both parties.

[14] Earlier versions of this paper used an alternating-offers bargaining model with off-path belief restrictions and derived an equivalent set of theoretical results. The somewhat artificial nature of the current wage-setting structure dramatically simplifies the presentation without fundamentally changing the results.

cially, this assumption guarantees that, once we endogenize the job-finding rate, disparities in new match formation rates are solely caused by lower demand for black workers, and not the bargaining process.[15] Fortuitously, the assumption also results in simple solutions.

Jointly, our assumptions will ensure that every match will find a mutually acceptable wage, that equilibrium in steady-state will be unique, that wages are uniformly distributed over the contract curve, and that they are on average equal to the equal-weights Nash bargaining solution (between a firm with beliefs given by θ and an α worker), despite the asymmetric information.

2.5 Steady State

A steady state of a labor market is a mass of α job seekers, a mass of β job seekers and a mass of monitored β workers along with equilibrium firm and worker wage acceptance and monitoring strategies that make these populations constant over time. There are three kinds of steady states: those in which all employees are monitored until match quality is revealed, those in which no monitoring occurs, and those in which the monitoring choice depends on the wage draw. As the latter are unstable to small perturbations in the parameters, we focus on the former two.

Consider the case where no employees are monitored: the white labor market. Matches never deteriorate and therefore the only source of job seekers is newly born workers. In this scenario, a firm just matched with a worker infers his probability of being of type α is the population prevalence g; the chance of a white job-seeker being good at a job to which he is matched is therefore

$$\theta_W = \theta_0 = g + (1-g)\beta.$$

Now suppose that all newly hired black employees are monitored and all bad matches are terminated. Newly matched black workers will be worse than average.

Lemma 1 *The probability a newly hired black worker is in a good match is*

$$\theta_B = \frac{\beta}{\beta g + (1-g)} < \theta_W. \tag{2}$$

Proof. See A.1 ■

[15] This of course makes F an equilibrium object, as the acceptability of wages in turn depends on F; but the solution is unique given our myopia assumption. We could instead assume F is uniform on $[0, q]$ but then unacceptable wages would be encountered, and the probability of disagreeable wages would vary between matches with black and white workers. Notice that F does not depend on worker type as due to assumption (C1) there are no wages the firm and β workers would accept but α workers would not. As equilibrium acceptable wages will form an interval, we could, instead of a uniform, use any distribution with connected, compact support by scaling it to the acceptable wage interval.

Therefore, although monitoring may be individually prudent for each firm, it creates a negative externality by feeding a stream of workers who are worse than the population average (i.e. containing more β types) back into the job-seeker pool. Surprisingly, the steady state θ_B of this process does not depend on the rate of information λ, the worker matching rate μ, or the rate at which new workers enter the market.[16]

2.6 Parametric Assumptions

Now we impose certain restrictions on the joint values of parameters sufficient to ensure the existence of both steady states.

For an equilibrium with no monitoring to exist for white workers, we want to assume that monitoring costs are not too low. For monitoring not to be optimal, the instantaneous monitoring cost must not be worth paying to detect bad matches, accounting for the fact that the cost must be recouped on the surviving fraction of workers.

$$\text{(C2)} \qquad \underbrace{\frac{b}{\lambda}}_{\text{Monitoring cost}} > \underbrace{(1-\theta_W)\frac{\lambda c}{r}}_{\text{Reduction in losses to errors}} \cdot \underbrace{\theta_W}_{\text{Proportion of remaining workers}}$$

Our second condition, antisymmetrically to (C2), posits that "monitoring costs must not be too high"and ensures that all new black employees will be monitored in equilibrium. As the monitoring decision is increasing in the wage, for there to be no monitoring in the black labor market, a condition is needed at the lowest wage in that market. As the lowest wage in the market depends on the speed at which workers match rather than simply the firm's break-even point, the relevant expression is a bit more complex.

$$\text{(C3)} \qquad \frac{b}{\lambda} < (1-\theta_B)\frac{\lambda c\,(\theta_B\mu + 2r) - 2rq}{r\,(\mu + 2r)}$$

In other words, $\theta_{B,}$ the belief about the average ability in the black unemployed pool, must be sufficiently low that acceptably high wages are only possible if the firm monitors. Strictness of the inequality ensures that switching to an unchurned market is not simply a matter of switching equilibria (as the non-monitoring one will not exist here). Notice that as the matching frictions vanish (μ increases), (C3) becomes the opposite inequality of (C2) for θ_B rather than θ_W.

Finally, for both labor markets to exist, it must be that workers can be in expectation gainfully employed; a sufficient condition is that the expected non-monitoring product of

[16] This is an artifact of the assumption that workers are infinitely lived.

workers drawn from the black unemployed pool is positive:

$$\text{(C4)} \qquad q - \lambda c(1-\theta_B) > 0.$$

3 Solution

First, we will use the model's properties to characterize the firm's and worker's actions. The main intuition behind the following result is that the firm is more willing to monitor if the bad matches terminated by monitoring are costlier, due to higher wages.

Lemma 2 *The firm's monitoring decision is increasing in the wage w and decreasing in its belief about match quality θ.*

Proof. See A.2 ■

Our next result shows that the wages acceptable to both the firm and the workers form an interval.

Lemma 3 *For a labor market with expected match quality θ, there is an interval of wages, $[\underline{w}_\theta, \overline{w}_\theta]$, the worker and firm both accept.*[17]

Proof. see A.3 ■

An intervalic structure for the wages in each market will simplify analysis significantly. From Lemma 3 we have that the mutually acceptable wages are an interval $[rU^\alpha_\theta, \overline{w}_\theta] = [\underline{w}_\theta, \overline{w}_\theta]$. As α workers never separate once they find a job, we have that the lowest wage is equal to the expected wage they'd get at another firm, adjusted for search time: $\underline{w}_\theta = \frac{\mu}{\mu+r}\int_0^q w dF = \frac{\mu}{\mu+r}[.5\underline{w}_\theta + .5\overline{w}_\theta]$, so that

$$\underline{w}_\theta = \frac{\mu}{\mu+2r}\overline{w}_\theta. \tag{3}$$

We now present the main theoretical results of the paper: existence and uniqueness of equilibria in the two markets that perpetuate their associated steady states.

3.1 The Non-Monitored Market

Proposition 1 *Assuming (C1)-(C4), the white (non-churned) labor market has a unique solution where the monitoring technology is not used. The average wage in this market is*

$$w^{avg}_{\theta_W} = \frac{\mu+r}{\mu+2r}\left[q - (1-\theta_W)\lambda c\right]. \tag{4}$$

[17]Incentives are weak at the interval's endpoints, but this is immaterial as F will put zero probability on them.

Proof. see A.4 ■

The main intuition for the proposition comes from Lemma 2. Since the value of monitoring is increasing in w, for a non-monitoring solution we need only check whether the firm chooses to monitor at the break-even wage $\overline{w}_{\theta_W}$. And (C2) ensures that monitoring does not occur at that wage.

Interestingly, since the firm cannot learn the worker's type in this non-churned equilibrium, type has no effect on wages.

3.2 The Monitored Market

Here, as workers are monitored, β workers sometimes face separation and therefore have a low outside option. However, they cannot accept low wages at which monitoring would not occur at beliefs θ_B without revealing their type; thus, such wages are not accepted by the firm. Therefore this equilibrium is effectively a pooling one as well, despite the fact β workers receive significantly lower utility than α workers.

Proposition 2 *Assuming (C1)-(C4), the black (churned) labor market has a solution where the monitoring technology is used in every match. The average wage in this market is*

$$w_{\theta_B}^{avg} = \frac{\mu + r}{\mu + 2r}\left[q - \frac{r(\lambda c(1-\theta_B) + b)}{\lambda\theta_B + r}\right]. \tag{5}$$

Proof. see A.5 ■

The intuition here again comes from (2), which tells us that the monitoring decision is increasing in w and therefore if monitoring occurs at $\underline{w}_\theta$ it occurs at all matches, and (C3) which ensures this condition holds. As the equilibrium strategies induce monitoring at every equilibrium wage, employees who are revealed to be in bad matches separate from the firm. This sends only β workers back into the job-seeking pool, churning the market quality to θ_B.

4 Implications for Labor Markets

The previous sections establish conditions under which there are two distinct steady-states of the labor market. In this section, we compare labor market outcomes for workers in these steady states. We first discuss a prediction that has not previously been tested and then discuss the relation of our other predictions to known labor market regularities.

4.1 Job Duration

Absent monitoring, there is no new information to dissolve the match. Therefore, taken literally, the model implies no turnover in the white equilibrium. In contrast, with monitoring, some workers prove ill-suited for the job and return to the unemployment pool. We interpret this as predicting that black workers will have lower average employment duration. Recall that workers who return to the unemployment pool are all type β. Therefore, turnover is even higher than if only new entrants were monitored. The model, again taken literally, implies that the separation hazard for black workers is

$$h_t = \frac{(1-\beta)(1-g)\lambda e^{-\lambda t}}{1-(1-\beta)(1-g)e^{-\lambda t}} \tag{6}$$

which is decreasing in t.

Importantly, h declines with t and asymptotes to 0, the hazard rate for whites. We expect this prediction to be robust to consideration of important real world elements not addressed by the model. Whether the hazard rates, in fact, converge is not something we are aware of the literature addressing and is the subject of our empirical investigation later in this paper.

As our model abstracts from firm-to-firm hiring, we have no prediction with regard to it. Although it may seem that firms would be out to poach black workers with high seniority (that are likely to have passed monitoring), adverse selection effects (with the worst workers more willing to leave) could unravel such effects, depending on the ability of outside employers to commit. Still, our predictions are in terms of employer-initiated separations, not moves to better jobs. Therefore, in the empirical section, we treat spells that end in a job-to-job transition as censored and, in the main specification, treat all quits as censored.

4.2 Ability, Race and Job Duration

We take the prediction in the last subsection one step further. Taken literally, our model implies that white workers are never laid-off regardless of their type. In contrast, black βs but not black αs are sometimes laid-off depending on whether they are good matches. We interpret this as a prediction that black layoffs will be more responsive than white layoffs to a measure of unobserved worker quality.

We note that this prediction stands in sharp contrast with previous interpretations of statistical discrimination. Thus Altonji and Pierret (2001) argue that statistical discrimination implies a smaller effect of AFQT on wages for black than for white workers, at least at low tenure. Arcidiacono, Bayer and Hizmo (2010) find no difference in the effect of AFQT on wages for black and white high school graduates and a negligible effect for all college graduates.

4.3 Wages

As $w_{\theta_B}^{avg} < w_{\theta_W}^{avg}$, black workers, on average, earn less than white ones. The highest wage firms are willing to pay is lower for black than for white workers since the average quality of new hires is lower, and the lowest wage black workers are willing to accept is lower because they expect other employers to pay less, as well. Interestingly, because white workers are not monitored, their lifetime utility does not depend on type, and both types have higher utility than black αs who, in turn, have higher lifetime utility than black βs:

$$U_{\theta_W}^{\alpha} = U_{\theta_W}^{\beta} > U_{\theta_B}^{\alpha} > U_{\theta_B}^{\beta}.$$

It is less obvious whether the wage distributions of black and white workers will overlap, and for some parameter values they do not. Figure 2 illustrates an example of the model where their equilibrium wage distributions significantly overlap, despite the fact only the latter are monitored.

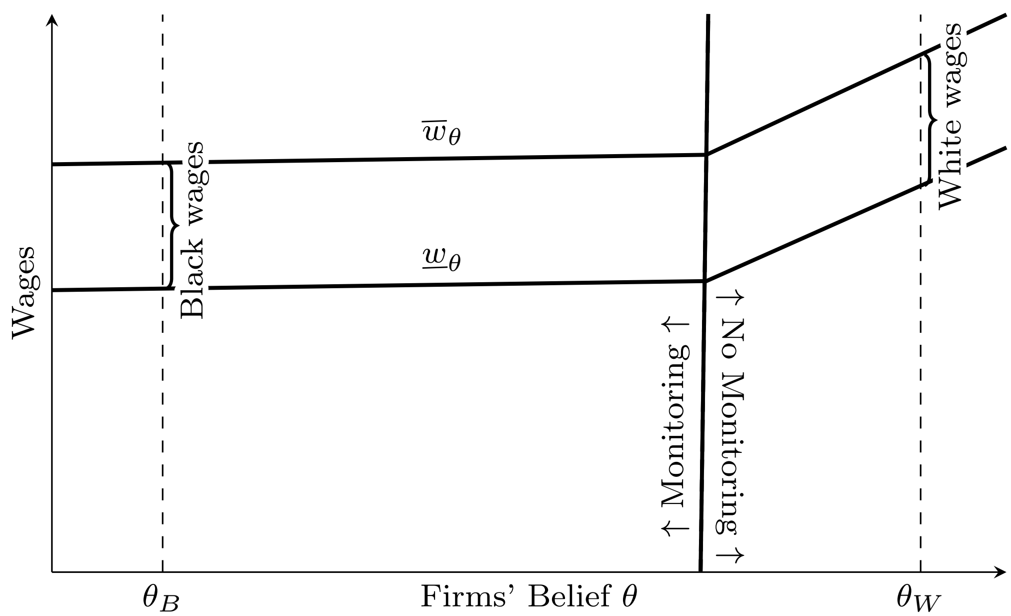


Figure 2: An example where the black and white wage ranges overlap. ($r = .05,\ g = .99,\ \mu = 2,\ \beta = .1,\ q = 1,\ b = 2,\ \lambda = 2,\ c = 4$)

4.4 Persistence of Discrimination

A key result of the churning mechanism in this paper is that deleterious steady states are persistent. In this section we show just how hard it is to transition to a good steady state. We regard this as illustrating the difficulty of addressing labor market discrimination in the context of policy, particularly policy aimed at improving the skills of black workers. The existence of a range of g values for which both steady states exist allows us to talk about *persistence* of the deleterious equilibrium.

Heretofore we have assumed that average skill levels for the two population groups are identical. Suppose instead that skill levels are $g_B \neq g_W$ and that the steady-state of each market provides monitoring for black but not white workers. Monitoring will persist as the equilibrium in the black labor market until g_B rises above some critical level while the no monitoring equilibrium will persist in the white market provided that g_W remains above a lower critical level. In principle, we can have the black workers in the bad equilibrium and the white workers in the good equilibrium provided that (C2) and (C3) hold for θ_W and θ_B calculated using g_W and g_B respectively. Put simply, this means that discrimination in wages and monitoring (and therefore also separations) can continue even if black workers are significantly better, on average, than white workers.

4.5 Unemployment Duration

We have so far treated the workers' matching rate, μ, as exogenous. Making the standard assumption of free entry, we now allow firms to post and maintain vacancies at a cost k per unit time. When a firm creates a vacancy, it can direct its search. This can take several forms, most notably locating production operations in an area with specific population characteristics or advertising the vacancy in different areas and through different media. In general, a firm can target markets indexed by i where a proportion ρ_i of unemployed workers are white. The open vacancy cost k is invariant to this target choice. We assume that in each market i the bargaining equilibria and population group steady states break down along the discriminatory lines described so far.

Define ϕ as market tightness and let the worker job-finding rate function follow the commonly assumed form

$$\mu(\phi) = m\phi^{\gamma} \tag{7}$$

for constants $m > 0$ and $0 < \gamma < 1$. Note that if firms expect a match to be worth V, the free-entry level of ϕ in such a market sets

$$\frac{\mu(\phi)}{\phi}V - k = 0 \tag{8}$$

so that

$$\phi = (\frac{Vm}{k})^{\frac{1}{1-\gamma}}. \tag{9}$$

Therefore, ϕ is an increasing function of V.

Assuming the parametric assumptions hold for the entire breadth of derived matching rates, we can now derive the free-entry equilibrium level of μ_{ρ_i} for each market i. The payoff to a firm for matching is the same as for an α worker, that is, when hiring from pool i, the firm expects a successful match to pay

$$V_i = \rho_i \frac{1}{\mu + 2r}\left[q - \lambda c(1-\theta_W)\right] + (1-\rho_i)\frac{1}{\mu + 2r}\left[q - \frac{r(\lambda c(1-\theta_B)+b)}{\lambda\theta_B + r}\right] \tag{10}$$

The above expression is increasing in ρ_i. Therefore, for the same μ, markets with more black workers will have a lower expected payoff for a filled vacancy. Therefore, the free-entry $\phi(\rho_i)$ and $\mu(\phi(\rho_i))$ are increasing in ρ_i, so that workers searching for jobs in markets with a higher proportion of black workers take longer, on average, to find employment. A first-order stochastic dominance argument can then show that black workers take longer, on average, to find employment, and receive lower average wages.

4.6 Extensions

4.6.1 Eventual revelation in all matches

We have assumed unrealistically that the match quality of workers who are not monitored is never revealed. More plausibly, heightened scrutiny speeds the rate at which match quality is revealed. In a model in which workers live forever, this change considered in isolation would eliminate our results because the composition of the jobless pool is independent of the rate at which bad matches are revealed. However, if workers do not live forever, then reducing the rate at which match quality is revealed does affect the quality of the unemployment pool, and our basic results go through.

4.6.2 Skill level and discrimination

Further, we can allow for observable heterogeneity among workers. If there are groups of workers for whom g is high, only the no-monitoring equilibrium will exist for these groups, regardless of race. This is also true at very low g and very low β (although we have assumed away this case to simplify the proofs). The first result is consistent with similar outcomes for black and white workers with high levels of skill as measured by education or the Armed Forces Qualifying Test (Neal and Johnson, 1996; Lang and Manove, 2011). The latter is consistent with some evidence that the bottom of the labor market is similarly bad for black and white workers. On the other hand, Lang and Manove find that the market learns the productivity of white but not black high school dropouts. This is consistent with an equilibrium in which white unemployed dropouts are, on average, more skilled than black unemployed dropouts and therefore in which white but not black dropouts are monitored. Nevertheless, without additional, largely *ad hoc* assumptions, this story cannot account for the very high unemployment rate among black dropouts.

4.6.3 Investment in unobservable skills

We have heretofore postulated that the proportion of α types is exogenous. Assume instead that some fraction of workers are innately of type α. Others can transform themselves from βs into αs at some cost ω with cdf $F(\omega)$. Provided that the fraction of natural αs satisfies (C2) and (C3), both equilibria will continue to exist for the right choice of F. However, since in the no-monitoring equilibrium αs and βs receive the same wage, there is no incentive to invest in becoming an α. In contrast, in the monitoring equilibrium, lifetime earnings are strictly higher for αs than for βs. Thus, some individuals will have an incentive to make the investment.[18] This prediction contrasts with Lundberg and Startz (1983) and Coate and Loury (1993), where black workers are less willing to invest in skills.

4.6.4 Education

Suppose now that there exists a signal,[19] which we identify with education, that α workers can purchase at some personal cost $\kappa \sim F$. Assume doing so ensures that any employer will be immediately aware that the worker is indeed type α. An educated α worker of either population with cost κ will then anticipate a lifetime utility of $V_{Educ}(\kappa) = \frac{\mu}{\mu+2r}q - \kappa$. In Section 4.3 we showed white α workers receive a higher lifetime payoff than their black

[18]It might appear that the incentive to undertake such investments would unravel the monitored equilibrium. However, if this were the case, no worker would have an incentive to invest. This raises messy dynamic issues which we sidestep by assuming that the fraction of additional workers who would choose to invest is insufficient to overturn (C3).

[19]We analyze the case of a pure signal. If education can also turn a β into an α, the analysis is a combination of the analysis in this and the prior subsection since productive investment increases the fraction of workers who are α but investment that reveals workers to be α reduces the fraction of unrevealed workers who are α.

counterparts; therefore, the incentive for the latter to invest in education is greater. As this implies that $\bar{\kappa}_W \equiv max\{\kappa : V_{Educ}(\kappa) \geq U^{\alpha}_{\theta_W}\} < max\{\kappa : V_{Educ}(\kappa) \geq U^{\alpha}_{\theta_B}\} \equiv \bar{\kappa}_B$, we must have that $F(\bar{\kappa}_W) < F(\bar{\kappa}_B)$ and therefore more black workers will purchase education. In particular, there exists some range of idiosyncratic costs for which black workers will purchase education but white workers will not. This is consistent with the finding in Lang and Manove (2011) that, conditional on past test scores, black workers get more education than white ones do. The intuition here is simple; if a worker of high skill is treated as if she has the average hire's skill for her group, she has a greater incentive to reveal her high skill if that average is lower.[20]

Perhaps equally importantly, this extension suggests that black and white workers with high observable skills will have similar outcomes as discussed in subsection 5.1.

4.6.5 Imperfect monitoring

We show in Cavounidis and Lang (2015) that one can write a very similar model in which β workers always match badly but monitoring can result in false positive good matches. Much of the analysis would remain unchanged in such a model. Parameters would exist that would force monitoring on black but not white workers, the black labor market would churn, and it would produce lower employment duration, higher unemployment duration and lower lifetime wages for black workers. In this formulation, black workers succeeding at monitoring would only be as good as whites who had never been monitored; therefore a churned market does not necessarily produce better long-run matches for the successfully monitored.

However, this alternate model would imply that some workers are purely parasitic and cannot be matched well, but rather aspire simply to find a job where their lack of productivity is undiscovered.

4.6.6 Stigma and degeneration into lower-skilled jobs

Our model assumes unrealistically that employers have no information regarding the time that workers have been in the labor market or the number of jobs they have held. If the other aspects of our model were a rough representation of reality, it is implausible that firms would not recognize that some workers were unlikely to be new entrants and therefore very likely to be β types. Then workers who did not find a good match sufficiently quickly would be permanently barred from the monitoring sector.

Somewhat more formally, as an extension to the model, we can relax the assumption that past history is entirely unobservable. Assume instead that each separation has a probability

[20] Strictly speaking, this creates a feedback loop from lower wages for the uneducated to a greater measure of education. The right assumptions on F rule out associated complexities.

ζ of becoming public common knowledge. Any worker who has a revealed separation is known to be of type β in any new match. Thus, a newly hired worker who does not have such a stigma will be of average quality $\theta'_B = [\beta + g\zeta(1-\beta)] \,/\, [g\beta + (1-g) + g\zeta(1-\beta)]$. If we assume θ'_B satisfies (C3), churning can persist but will be primarily a phenomenon for relatively young workers.

But what will happen to workers revealed to be βs? It is straightforward to extend the model to allow for a second occupation type (q', c') lacking monitoring technology[21] that is less skill intensive than the task described so far, i.e. $q > q'$ and $q - \lambda c < q' - \lambda c'$. Unrevealed β types can be strictly better off than revealed ones in a new match of the first task, but revealed βs are forced to enter the market for the second occupation.

In this scenario, a fraction of black workers are relegated to low-wage jobs while white workers with similar skills can always get better jobs.

4.6.7 Changing screening and monitoring technology

Autor and Scarborough (2008) examine the effect of bringing in a new screening process. They find that the screening process raised the employment duration of both black and white workers with no noticeable effect on minority hiring. In our model, we can think of this technology as allowing the firm to screen for job match quality prior to employment. This increases the proportion of hired black workers who become permanent since some bad matches are not hired. If the screening mechanism is good enough, the firm will choose not to monitor the black workers it hires, and all black workers will be permanent. Formally, since all white workers are permanent in the absence of the screen, the screen does not affect this proportion. Informally, if poor matches are more likely to depart even without monitoring, then there will also be positive effects on white employment duration.[22] Similarly, Wozniak (2015) shows that drug testing increases black employment and reduces the wage gap; we interpret this as confirming evidence for the notion that employers are more uncertain about the quality of black workers, and therefore that black workers benefit more from early resolution of such uncertainty.[23]

We note that improved technology appears to have reduced monitoring costs. This is unambiguously good for black workers who share the cost of being monitored. Unless the reduction shifts whites into the monitoring equilibrium, they are unaffected by the cost

[21] Or, more palatably, the same technology but without the incentives to use it, as in the case of a small enough c'.

[22] Formally, the model would have to be modified to ensure that some β workers are never perfectly matched and/or that some β workers are still in bad matches when they exit the labor force.

[23] Wozniak (2015) is not to be interpreted as evidence that monitoring is good for black workers on the aggregate. As in the present paper, it can be beneficial on an individual level (as it allows good workers to get higher wages than otherwise); our model, however, shows it can also create a worse externality.

reduction. However, if firms begin monitoring, α workers and firms will initially be better off. Firms will be able to better screen their workers, and as a consequence can offer higher wages, which should make α workers better off as the monitoring does not put them at risk. On the other hand, β workers will generally be worse off. In a collective bargaining setting, the union might resist monitoring. The more interesting point is that since monitoring creates an externality, it is easy to develop an example in which monitoring makes both types of workers and capital worse off in the long run.[24]

4.6.8 Job-to-job transitions

There are many ways to incorporate job-to-job transitions in a search model. Some would not alter our results qualitatively, and others would, while yet others would produce a multiplicity of equilibria. Of particular import is the information accessible to outside firms. Under the simplest assumption, that of information symmetry across firms, an outside firm that meets an employed worker learns whether monitoring has concluded successfully (or, alternatively, the worker could disclose this). Keeping with the wage offer structure, such an outside firm and worker pair would draw a wage uniformly among all mutually acceptable wages.[25] Wages would jump on transition, but our predictions regarding average wages by race would be preserved, conditional on any of age, experience, or job spell duration. The monitoring state of the worker would persist from firm to firm (as incentives to monitor increase in the wage, but white workers are not monitored even at firms' break-even wage) so that our predictions regarding employment spell duration go through.

5 Empirics

5.1 Are Black Workers Monitored More?

Our theoretical analysis assumes that black and white workers are in different equilibria and that, consequently, black workers are more heavily monitored. Of course, as we discuss in section 4.6.6, it is possible that some black workers end up in low-wage jobs where monitoring is unnecessary, after employers observe their previous separation history. In addition, a more

[24]Suppose that g_0 is just sufficient to sustain a no-monitoring equilibrium. A small reduction in b puts the labor market into a monitoring equilibrium. Initially, α workers and firms would experience a slight gain, but the churning will wipe this out and more. Firms always make zero profit on vacancies, but if we allow for a distribution of vacancy costs, then the rents earned by firms with low costs of creating vacancies will also fall.

[25]A stronger myopia assumption would be sufficient to retain our solution without additional parameter constraints.

realistic model would have black workers being more heavily monitored than white workers but would not predict that white workers are never monitored.

However, the spirit of the model is that black workers are monitored more heavily, and therefore, we look for direct evidence on the relation between race and monitoring. The evidence we have been able to find is very limited. We rely on a single question from the 1977 wave of the Panel Survey of Income Dynamics (PSID) that asks whether the respondent's supervisor checks his work "several times a day, once a day, once a week, every few weeks, or less often than that." We code the reported level of supervision from one to six, where six denotes supervision several times per day, and one denotes reporting not having a supervisor.

We do not know the nature of the supervision. Ideally, we would like a measure of supervision designed to assess the worker's quality, rather than supervision to prevent shirking or other malfeasance. Neal (1993) uses this variable to study differences in supervision focused on the latter. Still this is what appears to be available to us.

We restrict the sample to household heads who are actively employed or temporarily laid off private-sector workers, and who are not themselves supervisors.[26] We estimate the model by ordered probit and weight by the 1977 family weight. Early experimentation found only weak evidence of monitoring differences when we did not further restrict the sample and no evidence that black workers with more than a high school education were monitored more frequently than their white counterparts. As discussed in the theory section, our model need not apply to more educated workers who have signalled their higher productivity.Alternatively, the question may not be good at revealing monitoring differences among more skilled workers. Nevertheless, in the remainder of the paper we focus on workers with no more than a high school education.[27] Recall that black workers with more than a high school education were a relatively elite group during this time period. Only 11% of the black workers in the sample (16% weighted) were in this group.

The first column of Table 1 presents the results with no controls; black workers are more likely to report that they are monitored frequently. However, the coefficient is imprecisely estimated and significant at only the relaxed .1 level.[28] Including twelve occupation and

[26]We also restrict to respondents living in the United States who report a wage. Respondents are only asked to report a wage if they report being salaried or being paid hourly. Respondents who replied "other"or "NA; Don't Know"to whether they were paid by the hour or salaried were not asked for their wage. The question about supervision is also asked separately to individuals who report working for someone else and being self employed. We do not include these individuals as we cannot separately identify the occupation and industry for these two jobs. We exclude supervisors, as black workers who are supervising other workers are intuitively more likely to have passed monitoring.

[27]Even if earlier versions of this paper were not in circulation, honesty would require that we admit that our initial approach used workers regardless of education. In general, our results are strengthened by restricting the sample, but we did not think to use this restriction until we began to look at supervision directly.

[28]Throughout this subsection, we use one-tailed tests because we will not consider large negative t-statistics

Table 1: Likelihood of Employer Monitoring by Race

Y = Level of Supervision	(1)	(2)	(3)	(4)
Black	0.150	0.212	0.176	0.185
	(0.099)	(0.104)	(0.110)	(0.110)
Other Race	-0.248	-0.225	-0.350	-0.334
	(0.176)	(0.199)	(0.208)	(0.211)
Completed Education	≤ 12	≤ 12	≤ 12	≤ 12
Occupation, Industry FE	N	Y	Y	Y
Other controls	N	N	Y	Y
Ln(Hourly Wage)	N	N	N	Y
N	1,095	1,095	1,089	1,089

Notes: Robust standard errors in parentheses. Estimates are from an ordered probit using data from the 1977 PSID. Dependent variable: level of employer supervision. A value of six corresponds to employer checking the individual's work several times per day, five to once a day, four to once a week, three to every few weeks, two to less often, and one corresponds to no supervisor. Other controls: highest grade completed, age, age squared, tenure, tenure squared and indicators for temporarily laid off, south, north central, northeast, salaried, male, and union job. The sample includes household heads employed or temporarily laid off by private employers, who reported a wage and are not themselves supervisors. Observations are weighted by the family weights of the survey. See text for details.

eleven industry fixed effects (column 2, see Appendix B.1 for details) increases the coefficient, which is now significant at the .05 level. Note that adding these controls may be necessary or problematic. In an extended version of our model, we anticipate that black workers would be more likely to be matched with jobs in which monitoring is relatively inexpensive and/or in which monitoring is unimportant, perhaps because there is little opportunity for variation in output. In the former case, controlling for occupation and industry would obscure black-white differences; in the latter, it would be necessary. Adding these controls is a conservative approach.[29]

Column 3 includes additional controls for years of education, quadratics in tenure (truncated at the 99th percentile or 444 months) and age, male, union job, whether the worker is salaried, living in the Northeast, North Central, or South of the U.S. (with West the omitted region), and whether the worker is temporarily laid off. Note that some of these controls are potentially endogenous in a fuller model. This yields a smaller coefficient that falls just short of significance at the .05 level. Finally, as a potentially better but obviously

as evidence in favor of our alternative hypothesis, and we are not focused on identifying evidence that rejects no difference against the alternative that black workers are monitored less than white ones.

[29]For the coefficients on the controls, see Appendix Table B.1

endogenous control for worker skill, column 4 adds the natural log of the individual's wage, and the coefficient is again significant at the .05 level.[30]

Marginal effects in column 4 suggest black workers are 6.2 percentage points more likely to be monitored several times per day relative to white workers. Black workers are less likely to a) report no supervisor (2.3 percentage points), b) be monitored less often than every few weeks (2.6 pp), c) be monitored every few weeks (.5 pp), and d) be monitored once a week (.7 pp). There is no difference in the likelihood of monitoring once a day.

Our model implies that monitoring should decline faster with tenure for black workers than for white workers. We tried including interactions between race and tenure and tenure squared. Unsurprisingly given the small sample, the results were uninformative.

5.2 Unemployment, Race and AFQT

Building on a literature starting with Farber and Gibbons (1996), we use AFQT to capture both unobservable and observable predictors of quality such as education. It is well known that black individuals score lower than white individuals on the AFQT and increasingly well known that, conditional on AFQT, black workers get more education than white workers. Thus, it would be surprising if black unemployed workers did not have lower scores than their white counterparts on the AFQT. Nevertheless, for completeness, we verify this expectation.

We use the National Longitudinal Survey of Youth 1979, a nationally representative sample of 12,686 individuals, 14-22 years old when first surveyed in 1979, with oversamples of black, Hispanic and poor white individuals. These individuals are surveyed annually through 1994, and biennially afterwards. We restrict ourselves to the black and white samples and eliminate the poor and military oversamples, and use the most recent scaling of the AFQT. Respondent's current labor force status was recorded only in the waves through 1998 and again in 2006. We drop the waves after 1998 and, in each wave, individuals who were neither employed nor unemployed. We use the survey week labor force status variable from the NLSY, which is based on a question regarding the respondent's main survey week activity. Dropping individuals neither employed nor unemployed implies we exclude individuals whose main survey week activity was school enrollment or who were otherwise out of the labor force.

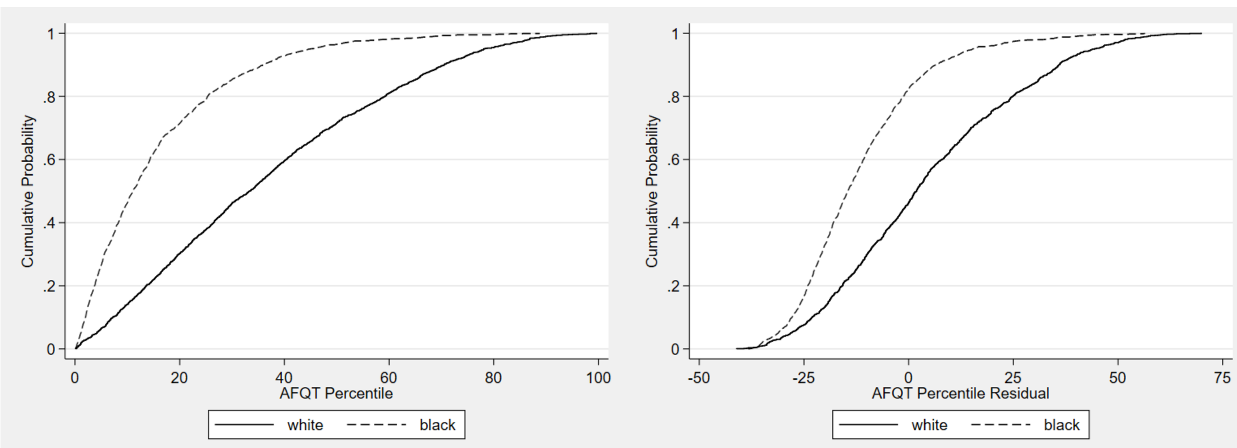
Absent first-order stochastic dominance (FOSD), whether one group has higher AFQT is scale dependent. Using the scale score rather than the percentile rank could change the ordering of groups. Consequently, we test for FOSD, or more precisely its absence, using the Kolmogorov-Smirnov test. Formally, this tests whether we can reject the null hypotheses

[30] Our model implies that white workers earn more than black workers. Thus, conditional on wage, black workers should be more skilled than white workers on average. If more skilled workers are less likely to be supervised, then controlling for wage should underestimate the true difference.

that the distribution of group A dominates that of group B and that B dominates A. If we cannot reject either hypothesis, we conclude that we cannot reject that the distributions are equal. Importantly, rejecting that A dominates B but not that B dominates A does not allow us to *accept* the null that B dominates A. However, when combined with visual evidence suggesting stochastic dominance, we will conclude that B dominates A.

The Kolmogorov-Smirnov test requires dividing the sample into two groups such as black and white workers. When comparing employed and unemployed workers some sample members may be employed in some years and unemployed in others. Our primary analysis divides workers between those who were unemployed at the time of any interview and those who never report being unemployed, whom we informally call "always employed" although they may have been out of the labor force at some interview.[31] We define education as the highest educational attainment the respondent reports in any year. Only respondents who never report education beyond high school are retained.

Figure 3: Distribution of ever-unemployed workers' AFQT by race.



Left panel: CDF of AFQT percentile. Right panel: CDF of residual from regressing AFQT percentile on education, SMSA residency, age, and region.

Not surprisingly, ever unemployed white individuals have higher AFQT scores than ever unemployed black individuals both visually - seen in figure 3 - and statistically. We can reject that black scores dominate white scores at any conventional level while the test statistic for the opposite hypothesis is < 0.0005. We also residualized AFQT using a regression of AFQT on education, a dummy for whether the individual lived in an SMSA at the time the AFQT was administered, their age in 1979 and three region dummies. Among those ever

[31]In principle, we could use multiple observations per individual and split the sample by whether the individual was employed or unemployed in a given year. However, the Kolmogorov-Smirnov test would then not have a standard distribution.

experiencing unemployment on an interview date, white individuals have higher residualized AFQT scores than black individuals both visually and statistically.

5.3 Does AFQT Reduce Black or White Unemployment More?

Our model also implies that unemployed black workers should be more adversely selected than their white counterparts. Unfortunately, we are unable to determine whether this prediction is empirically valid. For the reasons discussed in Bond and Lang (2013), it turns out that which group is more adversely selected depends on our choice of scale.[32]

Instead, we test whether AFQT reduces black unemployment more than white unemployment, an equally good test of the 'churning' theory. Later, we focus on whether higher AFQT scores reduce the layoff hazard more for black workers than for white workers.

We begin by regressing a dummy for whether the individual is unemployed that year on AFQT percentile, black and black interacted with AFQT percentile. We also control for education, whether the individual lived in an SMSA, age and three region dummies, and cluster on the individual.

The first column of Table 2 shows the results of this estimation. As predicted by the model, an additional AFQT percentile reduces the probability of unemployment among black workers substantially more than among white workers. The effect on the former is roughly twice the effect on the latter and is statistically significant at the .01 level.

Table 2: Effect of AFQT on Probability of Unemployment, by Race

	AFQT Measure	
	Percentile*.01	Standard Normal Distribution
Black	0.14	0.10
	(0.01)	0.06
AFQT	-0.10	-0.03
	(0.01)	(0.004)
Black*AFQT	-0.10	-0.02
	(0.03)	(0.01)
Observations	50046	50046

Notes: Linear probability model. Dependent variable equals 1 if unemployed, 0 if employed. Other controls: living in SMSA, age, educational attainment, three region dummies. Standard errors are clustered by respondent.

However, the Bond and Lang critique of such comparisons still applies. While we have

[32] That is, effectively, our choice of mapping from AFQT score to ability.

not replicated our result for all possible monotonic transformations, we wish to illustrate the robustness of the result.

In the second column, we translate the AFQT percentile into a standard normal distribution by inverting the normal. Using this scale, a one standard deviation increase in AFQT reduces the probability of white unemployment by three percentage points, but the effect on black unemployment is two percentage points greater (significant at the .05 level).

This finding is not robust to all transformations of the AFQT scale. For instance, the coefficient on the interaction term is negative but insignificant if our measure is ln(AFQT Percentile), which puts a lot of importance on differences at the very bottom. It is significant at the .1 level if our measure is ln(AFQT Percentile +8); it turns significant at the .05 level if we add 13 or more to the percentile. It is also significant if we instead use exp(AFQT Percentile).

In conclusion, we can choose AFQT scales that eliminate this result. However, it appears that unless we put a great deal of weight on differences at the very extremes of the AFQT scale, the results are robust.

5.4 Initial Black-White Gaps in the Layoff Hazard, and Convergence Over Time

We now test the model's prediction that the layoff hazard is initially higher for black workers but converges to that for white workers. To our knowledge this prediction has not previously been tested. The model also predicts longer unemployment durations and lower lifetime incomes for black relative to white individuals. These are known to be strongly empirically supported (see Lang and Lehmann 2012).

5.4.1 Methods

We estimate the layoff hazard using the first full-time spell of each individual at each employer. We censor spells ending for any reason other than the employee being discharged or laid off. For these censored spells, we assert we do not know when the spell would have ended in a layoff. We use both nonparametric and semiparametric survival analysis methods to estimate the hazard.

First, using standard techniques, we calculate the hazard over time intervals with the intervals large enough not to require further smoothing. For each nonoverlapping time interval, $(t_{j-1}, t_j], j = 1...k+1$, we obtain the number of employment spells at the start, the number of spells ending in a layoff (failures) over the interval, and the number of spells ending but not in a layoff (censored). A conventional way of calculating the hazard in this setting is to assume that censoring and death times are uniformly distributed within each interval. The hazard at the midpoint m for each nonoverlapping interval is then :

$$\hat{h}(t_{mj}) = \frac{d_j}{[(t_j - t_{j-1})(Y_j - \frac{d_j}{2})]} \tag{11}$$

The variable Y_j is the number of spells at the start of the interval minus half of the spells censored over the interval, and d_j is the number of failures over the interval (Klein and Moeschberger 2003).[33]

We use intervals of 26 weeks through durations of 520 weeks. After this point, 26 week intervals no longer include at least one black worker who was laid off and at least one white worker who was laid off, and so we use intervals of 39 weeks. This facilitates comparison of black and white worker hazards in our subsequent Cox analysis, as these models provide estimates only at failure times.[34] We calculate the hazard separately over these intervals for black and white workers. We obtain confidence intervals based on the estimated standard deviation of the hazard function at the midpoint of interval j, using the property that the number of failures in the interval is a binomial random variable.[35]

This nonparametric method does not allow controlling for covariates, and so we additionally plot baseline hazard functions for black and white workers from a Cox proportional hazards model stratified by race. The stratified Cox model allows for different baseline hazard functions for black and white workers, rather than assuming their baseline hazards are proportional. As in the traditional Cox model, we constrain the coefficients on the covariates to be the same for black and white workers.

We use the time intervals defined above as our measures of time, so that the baseline estimates do not require further smoothing. Using time intervals rather than week as a measure of time creates more instances of failures occurring at the same "time", since time is now a larger unit.[36]

The baseline contributions we obtain from this model are the same as the Nelson-Aalen contributions in the case of no covariates, using the week intervals as a measure of time.[37]

[33] This is referred to in the literature as the life table method.

[34] These 39 week intervals include at least one black and one white worker failure through durations of 793 weeks (15.25 years). While we use all durations for estimation, our figures show hazard estimates through 793 weeks.

[35] The formula for the estimated standard deviation of the hazard is found in Klein and Moeschberger (2003) and used in STATA. A similar formula is derived in Gehan (1969).

[36] There are several methods for dealing with these ties, all requiring assumptions about the timing of these failures. We present results using the Breslow approximation, one of the conventional methods and the STATA default. This is based on the assumption that the subjects failed at different times, but we do not know the order.

[37] There are several estimators of the baseline hazard rate in a proportional hazards model. We use the estimator from Kalbfleisch and Prentice (2002), which is also the default in STATA.

Specifically, we estimate

$$h(t|W,\mathbf{Z}) = h_W(t)exp(\mathbf{Z}\gamma) \tag{12}$$

The variable W is an indicator for whether the individual is white and $\mathbf{Z}$ includes highest grade completed, indicators for geographic region (Northeast; North Central; South; West is omitted), whether the individual lived in an urban area, age, occupation and industry fixed effects, and year fixed effects, all measured at the start of the spell, and the AFQT percentile. As described in Appendix B.2.4, we classify occupation into 18 groups and industry into 14 groups.

While we are able to use the default STATA program for our main hazard estimates, there is no built-in program to obtain bootstrapped standard errors for the baseline hazard contributions and so we describe the estimation in detail. Following Kalbfleisch and Prentice (2002), we obtain the baseline hazard contributions at each failure time by maximizing the likelihood function

$$\Pi_{i=1}^{k}\left[\Pi_{j\in D_i}\left(1-\alpha_i^{\exp[Z_j(t_i)'\hat{\gamma}]}\right)\Pi_{l\in R(t_i)-D_i}\alpha_i^{\exp[Z_l(t_i)\hat{\gamma}]}\right] \tag{13}$$

where $1-\alpha_i$ is the baseline hazard at each failure time t_i, $i=1,...l$, D_i is the set of individuals who fail at time t_i, and $R(t_i)$ is the set of individuals at risk of failing just prior to time t_i. Maximizing (13) with respect to α_i implies the maximum likelihood estimate of α_i is the solution to

$$\sum_{j\in D_i} exp[Z_j(t_i)\hat{\gamma}][1-\alpha_i^{exp[Z_j(t_i)\hat{\gamma}]}]^{-1} = \sum_{l\in R(t_i)} exp[Z_l(t_i)\hat{\gamma}] \tag{14}$$

This is solved using an iterative procedure. Because these baseline contributions are functions of the estimated coefficients γ on the covariates, to obtain standard errors, we use a parametric bootstrap in which we draw 10,000 sets of coefficients using the variance-covariance matrix of our estimated coefficients and calculate the estimated baseline hazard contribution at each time using (14).[38] We use the standard deviation of these estimated contributions to form the confidence intervals.

For robustness, we determine the hazards at each week, rather than for an interval of weeks, and then smooth using a kernel-smoother and local linear smoothing. These methods require choices of kernels and bandwidths, and, in the former case, an approach to addressing bias in the boundary regions.

[38] For the iterative procedure to solve (14), we use the initial value suggested by Kalbfleisch and Prentice (2002): $1-\alpha_{i0} = d_i\{\sum_{l\in R(t_i)} exp[Z_l(t_i)\hat{\gamma}]\}^{-1}$, where d_i is the number of failures over the interval.

5.4.2 Data

We test the model's prediction using the NLSY79. We construct job spells using the Employer History Roster which greatly facilitates linking job spells across survey years, by assigning each job a unique identification number consistent across surveys.[39] We define job spells as the first full-time spell with each employer, defining full time as at least 30 hours per week. For each survey year in which an individual reported employment at a given employer, we collect the start and end week of employment with that employer reported in the survey. We construct the total length of the job spell by grouping all consecutive full-time spells at the employer across survey years. We treat gaps at the same employer of less than or equal to 26 weeks as continuations of the same job spell at the employer, but subtract the length of the gap from the duration.[40] We evaluate whether the individual is fired or laid off only at the end of the linked spell, and not at the time when the gap begins. This avoids treating very temporary layoffs similarly to more permanent discharges or layoffs. For robustness, we do not link noncontinuous reported spells at the same employer.

We focus on the layoff hazard, which we define as the hazard of a job ending due to the employee being discharged, fired, or laid off.[41] Job spells ending for other reasons are censored. For robustness, we treat quits into nonemployment similarly to layoffs.

Our sample includes non-Hispanic individuals who had obtained no more than a high school degree at the start of the job spell, consistent with our earlier analysis.[42] We exclude spells in which the worker ever reports self employment or working for a family business. We further exclude individuals with missing start or end weeks for any full-time spell, and individuals with full-time spells that end before they begin.[43] These missing or unclear start

[39] Information for jobs six through 10 reported in some of the early survey years may not have been added by NLSY to the roster due to difficulty recovering these data. This is unlikely to have a large impact on the results given these jobs are a small proportion of those ever reported, for a small proportion of individuals (National Longitudinal Survey of Youth 2019).

[40] The individual could report multiple job spells at the same employer, reporting a start and end week for each span. Additionally, for any given spell the individual can report a within-job gap at the employer and the start and end week of that gap. We exclude spells in which the individual reported a gap of more than 26 weeks, within the given start and end week they reported at the employer. The reported reasons for these within-job gaps make it difficult to identify whether the gap was due to the individual being fired or laid off, and so we exclude these spells.

[41] From 1979 through 1983, the NLSY groups together "layoff, plant closed, or end of temporary or seasonal job"as one reason for the job ending. Starting in 1984 these three categories are separated, and we treat layoffs as failures and the other two categories within the original group as censored.

[42] We show results for individuals with more than a high school degree at the spell's start for completeness. As above, we exclude the poor and military oversamples, and we include all surveys through 2010. We use the NLSY racial/ethnic cohort coding from the 1978 screener interview, which codes individuals as hispanic; black; or non-black, non-hispanic.

[43] Due to rounding week numbers, it is possible that the start week of the span is greater than the end week. In cases when the start week is up to two weeks greater than the end week, we replace the start week equal to the end week (National Longitudinal Survey of Youth 2019).

and end weeks make it difficult to know whether any of the individual's spells end in non-full-time employment, and our robustness analysis treats quits into nonemployment similarly to layoffs. In order to keep the samples similar, we exclude all spells for these individuals.

Because the survey is conducted every two years starting in 1994, we do not know the values of some of the control variables in some years and must impute their values from adjacent years. Appendix section B.2.1 describes these imputations in detail. In order to avoid excluding individuals with missing values of the covariates, we include an indicator for whether the individual is missing the value of the covariate, and set the value of the covariate to zero. As described in Appendix B.2.4, occupation and industry codes vary in the NLSY with the survey year. We use crosswalks from the Minnesota Population Center to convert all codes to 1990 Census occupation or industry codes.

As Table 3 shows there are nearly 34,000 job spells in the sample, which are the first full-time job spells at each employer for individuals in the sample. There are 20,140 job spells for white workers and nearly 13,700 job spells for black workers. For white workers the average spell duration is 91.5 weeks, while for black workers the average spell duration is 82.4 weeks, though this is underestimated due to censoring. On average there are over 5.7 job spells for each white worker in the sample, and 6.25 spells for each black worker in the sample. The proportion of job spells ending in a layoff or the employee being discharged is 21% for white workers and 23% for black workers. The table shows other differences between the average white and black job spells, including the worker's age, education, AFQT, urban location, occupation, and region. Importantly, the stratified Cox proportional hazard models will include these as covariates.

5.4.3 Results

Figure 4 shows the hazard estimates using bins. The patterns are consistent with our model. The layoff hazard for black workers is significantly higher than that for white workers starting in week 26, and continuing through 1.5 years of tenure, with nonoverlapping confidence intervals. After this, the gap decreases considerably, and the confidence intervals overlap. By 600 weeks of tenure (11.5 years), the point estimates suggest the layoff hazard for black workers is roughly equal or smaller than that for white workers.

Appendix Figure B2 shows nonparametric plots with hazards by week rather than larger bins, smoothed using kernel smoothers with various bandwidths and kernels, as well as local linear smoothing. Based on the absence of a gap in the first 26 weeks in figure 4, it is not surprising that due to smoothing most of these plots show a smaller or nonexistent gap in the first year followed by an opening of the gap. By year twelve, the point estimates suggest there is no gap in the layoff hazard.

Table 3: Summary Statistics

	White	Black
Spells Ends in Layoff/Discharge	0.21	0.23
	[.41]	[.42]
Spell Duration	91.5	82.4
	[178.8]	[158.7]
Male	0.55	0.61
	[.5]	[.49]
Age at Spell Start	25.8	27.3
	[8.01]	[7.9]
Highest Grade Completed at Spell Start	11.3	11.4
	[1.23]	[1.11]
AFQT (percentile)	42.3	18.6
	[25.2]	[17.5]
Urban Location at Spell Start	0.72	0.83
	[.45]	[.37]
Spells per Person	5.65	6.26
	[5.17]	[5.28]
Occupation: Managerial and Professional	0.06	0.04
	[.24]	[.19]
Occupation: Technical, Sales, Administrative	0.22	0.18
	[.41]	[.38]
Occupation: Service	0.18	0.25
	[.38]	[.43]
Occupation: Precision Production, Craft, and Repairers	0.13	0.1
	[.34]	[.3]
Occupation: Operatives and Laborers	0.23	0.29
	[.42]	[.45]
Spells at Risk of Ending in Non Employment at		
200 weeks	2298	1406
400 weeks	1071	578
600 weeks	587	288
800 weeks	331	160
1000 weeks	197	99
Total Spells	20140	13674

Notes: Standard deviations in brackets.

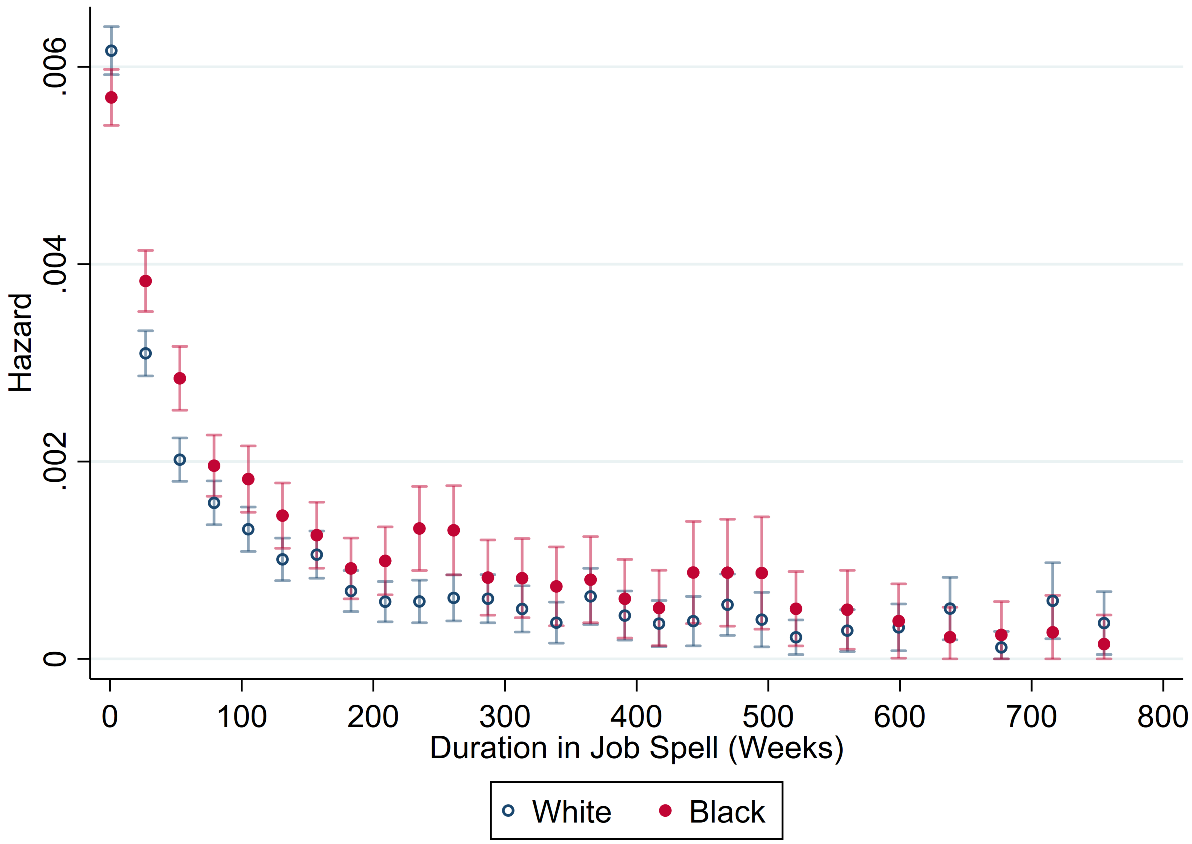


Figure 4: Nonparametric Estimates of the Layoff Hazard by Week Bins: first full-time spell at each employer. Confidence intervals based on the estimated standard deviation of the hazard function at the midpoint of the interval, using that the number of failures in the interval is a binomial random variable.

In Figure 5, we present Cox proportional hazard estimates using the multi-week intervals as units of time. The pattern is quite similar to that in Figure 4, though the point estimates suggest a similar, not a lower, hazard for black workers in the first 26 weeks, when adjusting for individual-level covariates.[44]

Appendix Figure B2 shows additional plots in which we estimate Cox regressions using week rather than larger bins, and smoothing hazard contributions using kernel smoothers with various bandwidth and kernels. Similar to figure 5, these plots show an early gap in the hazards of black and white workers, which becomes nonexistent by 12 years of tenure. Appendix Figure B2 further shows results using local linear smoothing of the hazard contributions, without any controls. These results are similar to the other nonparametric results

[44]See Appendix Table B2 for coefficients on the covariates. Omitting the region, industry, and occupation fixed effects in the Cox estimation yields a larger hazard for white workers than for black workers in the first 26 weeks.

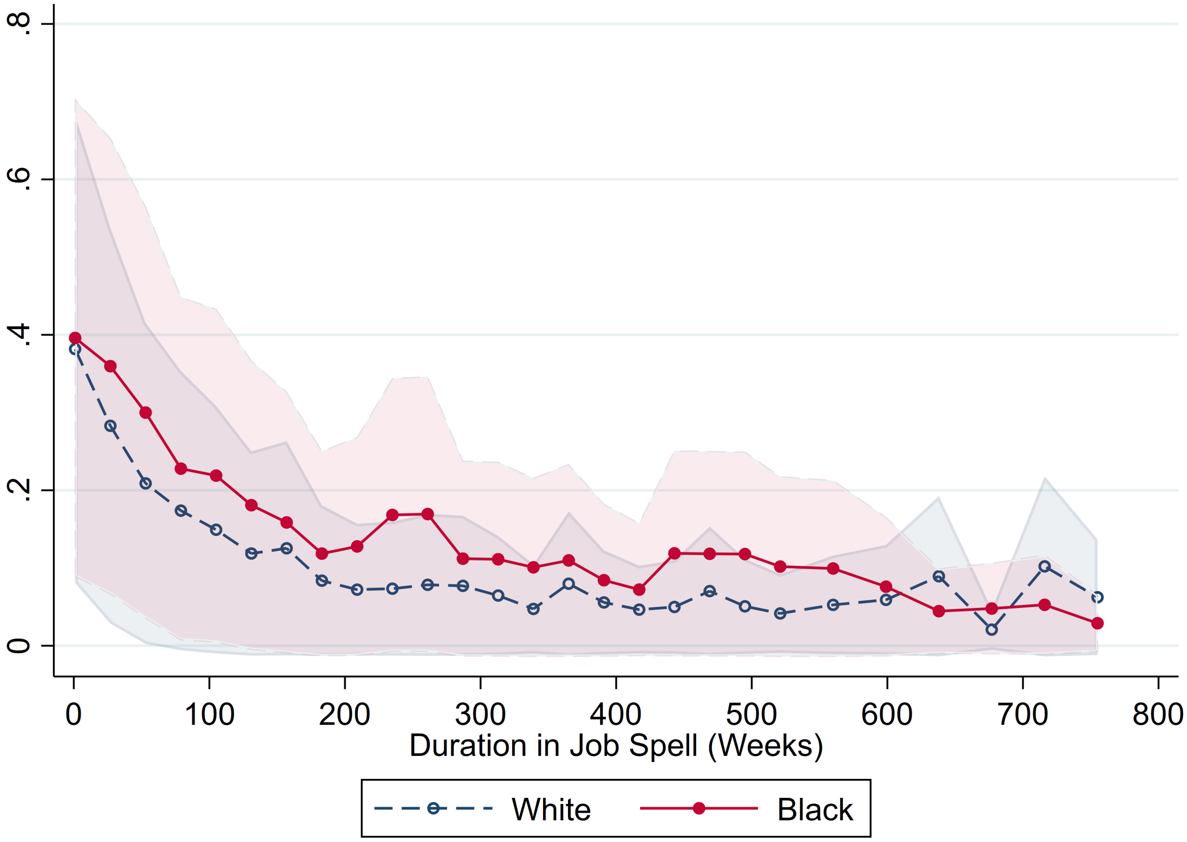


Figure 5: Estimates of the Layoff Hazard by Week Bins, Based on a Cox Model Stratified by Race. Confidence intervals based on a parametric bootstrap of 10,000 sets of coefficients. See text for list of included covariates.

in Appendix Figure B2, and show the hazards converge by roughly 12 years of tenure.

We also estimated a Cox model with the same covariates, but modeled the percentage gap between the black and white hazards to be a cubic in seniority. To allow the effect of race on the hazard to vary over time, we include an observation for each job spell at each failure time in the data (as Cox models are only estimated when failures occur). We find that the gap at week one is 3.3%, and this becomes statistically significant at the 10% level at week seven. Over the range from 1 to 793 weeks, the gap reaches its maximum of 64.5 percent at 266 weeks (roughly five years). It then falls, ceasing to be significant at the 5% level at approximately ten and a half years, and reaches zero at roughly 14 and a half years.[45]

[45] Modeling the percentage gap to be a quartic in seniority yields very similar results.

5.4.4 Robustness

The incentive to monitor new workers is relevant mainly for individuals hired out of nonemployment. For robustness, we identify employment spells for which the individual entered from nonemployment, and restrict to those spells.[46] We continue to see a gap emerging within the first year of the job spell, and convergence by the twelfth year (Appendix Figure B1(c) and (d)). As we discuss above, monitoring of newly hired black workers may be less likely among older individuals, for whom it is more clear that they are not new entrants to the labor force. Restricting to workers no more than 30 at the start of the spell, there is suggestive evidence that the gaps in the first two years are larger in magnitude, and remain nonoverlapping for an additional 26 weeks (Appendix Figure B1(g) and (h)).

In the main specification we treat gaps at the same employer of less than or equal to 26 weeks as continuations of the same job spell at the employer. For robustness, we treat any gap in employment at an employer as ending the spell, and evaluate whether the worker was fired or laid off at that point.[47] Appendix Figure B1(e) and (f) show the hazard functions for this robustness specification look quite similar to the principal results.

For robustness we treat quits into nonemployment similarly to layoffs.[48] We treat a job spell as ending in non-full-time employment if there is more than one week between the spell's end and the start of the next full-time job spell, reported at any other employer in any survey year. We alternatively treat a spell as ending in non-full-time employment only if there are more than four weeks between the spell's end and the start of the next full-time job spell. Both of these definitions yield similar results (Appendix Figure B1((a), (b), (m) and (n)).

The hazard of a job ending due to a plant closing is relatively flat for black and white workers, though this is underpowered with approximately 850 spells ending for this reason (Appendix Figure B1(k) and (l)). If these hazards looked similar to our main results this would raise concerns about other underlying mechanisms, as this is a setting without a role for monitoring differences.

When restricting to workers with more than a high school degree, there is much less evidence of differences in black-white hazards at early tenures (Appendix Figure B1(i) and

[46]We define an individual as being hired out of nonemployment if there is more than one week between the end of their last spell and the start of the current spell. Further, individuals are identified as being hired from nonemployment for their first full-time spell. If the individual did not respond to the next survey after the previous spell, and the year the last spell was reported was at least one year before the current spell was reported (two years earlier starting in 1996 when the previous survey was two years earlier), the individual is not coded as coming from nonemployment.

[47]This specification also excludes any spells for which the worker reported a gap within the start and end weeks of the span at the employer. This restriction excludes about 5100 spells.

[48]See Appendix B.2.3 for details on coding quits.

(j)). Similar to our findings above, this suggests the mechanisms in our model are most relevant for less-educated workers.

Table 4: Differential Effect of AFQT on the Layoff Hazard, by Race

	White	Black	All
AFQT (Percentile)	-0.0012	-0.0049	-0.0012
	(0.0007)	(0.0012)	(0.0007)
AFQT*Black			-0.0040
			(0.0013)
Black			-0.188
			(.222)
Observations	20,140	13,674	32,627

Notes: Conventional standard errors in parentheses. Coefficients are from a Cox Proportional Hazards model, using week bin as a unit of time. Each observation is a job spell. We model the layoff hazard, and the failure variable is an indicator for whether the job spell ended because the worker was fired or laid off. The regression additionally includes highest grade completed at spell start (and interacted with black in column 3), indicators for male, region at spell start (northeast, north central, south, and omitting west), urban location at spell start, age at spell start, fixed effects for year, occupation (18 groups), and industry (14 groups) all measured at the start of the spell, as well as indicators for whether AFQT, region, and urban location are missing. Column 3 excludes individuals with missing AFQT. In columns 1 and 2 we include an indicator for missing AFQT and replace AFQT with zero. See text for details.

5.4.5 Prediction: Layoff Hazard Declines with Ability for Black Workers More than for White Workers

We test one further model prediction related to ability and the layoff hazard. If ability reduces the likelihood of making mistakes, and monitoring allows employers to discover mistakes, then monitored workers who are lower ability should be more likely to be fired than monitored workers who are higher ability. This relation should be weaker for nonmonitored workers, as the mistakes of lower ability workers are less likely to be discovered. We test this prediction by estimating a separate Cox model for white and black workers, and comparing the coefficient on AFQT percentile, while recognizing the caveats based on the Bond and Lang (2013) critique discussed above.

There is a negative effect of AFQT on the layoff hazard, and the magnitude for black workers is four times the size of the coefficient for white workers (Table 4, for all coefficients see Appendix Table B3). The difference is significant at the 1% level. We also estimate a Cox model including both black and white workers, and an interaction between black and AFQT, as well as between black and highest grade completed given correlation between AFQT and education. This specification yields similar results, and the coefficient on the interaction is significant at the 1% level. These results present further evidence consistent with differential monitoring of black workers.

6 Conclusion

We develop a model that predicts known disparities between black and white workers: black workers earn lower wages, have longer unemployment duration, and obtain more education conditional on measured ability. It also predicts one previously unstudied disparity: the layoff hazard is higher for black workers at low tenure but the hazard rates converge as tenure increases. In addition, the effect of a measure of unobserved skills on layoffs should be more beneficial for black than for white workers.

As we have argued previously, while the model, of necessity, relies on some special assumptions, the key elements are 1) that worker productivity is correlated across jobs, 2) that ability is neither perfectly observed or signalled and workers can to some extent hide past firings, 3) that firms therefore use race to statistically infer worker ability, 4) that additional information arrives during employment and is either imperfect, costly or both so that a worker's productivity can never be known perfectly at zero cost, and 5) that firms can and do act on new information by firing some workers.

The predictions are largely confirmed. In our stratified Cox models, conditional on observables, at the beginning of a job black workers are more likely to be laid off than white workers. There is little evidence of a gap after roughly 10 years of seniority. We also confirm that higher unobservable skills, as measured by AFQT, more strongly reduces the likelihood that a black worker is laid off, relative to a white worker.

Contrary to the model's prediction, our results show there is no gap in the first half of a year on the job. Obviously, it is up to the reader to decide how problematic this is. Our interpretation is that this is a period when all worker/firm pairs are discovering whether they are grossly mismatched, information that both the worker and firm receive for free. The mechanism we underline becomes increasingly important with tenure and dominates after this initial probationary period.

Our message is in some ways depressing. Simply addressing education or human capital disparities between black people and white people need not eliminate labor market disparities. The 'bad equilibrium' in which many black people find themselves is difficult to escape.

References

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

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics*, 2 (4): 76-104.

Autor, David H. and David Scarborough, 2008. "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," *Quarterly Journal of Economics,* 123(1):219-277.

Becker, Gary S. 1971. *The Economics of Discrimination*, 2nd edition. Chicago, IL: Chicago University Press.

Bjerk, David. 2008. "Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Hiring and Promotion." *Economic Journal,* 118(530): 961-982.

Black, Dan A. 1995. "Discrimination in an Equilibrium Search Model." *Journal of Labor Economics*, 13(2): 309-333.

Bond, Timothy N. and Kevin Lang. 2013. "The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results." *Review of Economics and Statistics.* 95(5):1468–79/

Bowlus, Audra J., and Zvi Eckstein. 2002. "Discrimination and Skill Differences in an Equilibrium Search Model." *International Economic Review*, 43(2): 1309-1345.

Bowlus, A. J., Kiefer, N. M. and Neumann, G. R. 2001. "Equilibrium Search Models and the Transition from School to Work."*International Economic Review*, 42(2), 317-343.

Cameron, A. Colin and Pravin K. Trivedi. 2009. *Microeconometrics: Methods and Applications*, 8th edition. Cambridge, England: Cambridge University Press.

Cavounidis, Costas and Kevin Lang. 2015. "Discrimination and Worker Evaluation." *National Bureau of Economic Research Working Paper No. 21612.*

Charles, Kerwin, and Jonathan Guryan. 2008. "Prejudice and The Economics of Discrimination." *Journal of Political Economy*, 116(5): 773-809.

Coate, Stephen, and Glenn C. Loury. 1993. "Will affirmative-action Policies Eliminate Negative Stereotypes?" *American Economic Review,* 83(5): 1220–1240.

Coates, Ta-Nehisi. 2012. "Fear of a Black President." *The Atlantic.* Retrieved on January 12th, 2016 from: http://www.theatlantic.com/magazine/archive/2012/09/fear-of-a-black-president/309064/#

Epanechnikov, V. A. 1969. "Non-Parametric Estimation of a Multivariate Probability Density." *Theory of Probability and its Applications,* 14(1): 153-158.

Fairlie, Robert W. and William A. Sundstrom. 1999. "The Emergence, Persistence, and Recent Widening of the Racial Unemployment Gap." *Industrial and Labor Relations Review.* 52(2): 252-70.

Farber, Henry, and Gibbons, Robert. 1996 "Learning and Wage Dynamics." *Quarterly Journal of Economics*, 111(6): 1007-48.

Fryer Jr., Roland G. 2007. "Belief Flipping in a Dynamic Model of Statistical Discrimination." *Journal of Public Economics*, 91(5-6): 1151-1166.

Gehan, Edmund A. 1969. "Estimating Survival Functions From The Life Table." *Journal of Chronic Diseases*. 21: 629-644.

Guardian, 2014. "Prince: Black People Don't Get Second Chances," February 25, 2014, http://www.theguardian.com/music/2014/feb/25/prince-black-people-dont-get-second-chances, accessed January 24, 2015.

Kalbfleisch, J. D., and R. L. Prentice. 2002. *The Statistical Analysis of Failure Time Data.* 2nd ed. New York: Wiley.

Klein, J. P., and M. L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data.* 2nd ed. New York: Springer.

Lang, Kevin. 1986. "A Language Theory of Discrimination." *Quarterly Journal of Economics*, 101(2): 363-382.

Lang, Kevin and Jee-Yeon K. Lehmann. 2012. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature*. 50(4): 959-1006.

Lang, Kevin, and Michael Manove. 2003. "Wage Announcements with a Continuum of Types." *Annales d'Economie et de Statistique*. 71-72: 223-244.

Lang, Kevin, and Michael Manove. 2011. "Education and Labor Market Discrimination." *American Economic Review*, 101(4): 1467-1496.

Lang, Kevin, Michael Manove, and William T. Dickens. 2005. "Racial Discrimination in Markets with Announced Wages." *American Economic Review*, 95(4):1327-40.

Lauermann, Stephan and Wolinsky, Asher. 2016. "Search with adverse selection. "*Econometrica*, 84(1):243-315.

Lundberg, Shelly J., and Richard Startz. 1983. "Private Discrimination and Social Intervention in Competitive Labor Markets." *American Economic Review*, 73(3): 340-7.

Mabry, Marcus. 2007. *Twice as Good: Condoleezza Rice and Her Path to Power.* New York: Modern Times. Print.

MacLeod, W. Bentley, 2003. "Optimal Contracting with Subjective Evaluation." *American Economic Review,* 93(1):216-240.

Masters, Adrian. "Statistical discrimination from composition effects in the market for low-skilled workers." *Labour Economics,* 26 (2014): 72-80.

Moro, Andrea, and Peter Norman. 2004. "A General Equilibrium Model of Statistical Discrimination." *Journal of Economic Theory*, 114(1): 1-30.

National Longitudinal Survey of Youth. 2019. "NLSY 1979 Appendix 18: Work History Data. " https://www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/codebook-supplement/nlsy79-appendix-18-work-history-data, accessed, April 15, 2019.

Neal, Derek. 1993. "Supervision and Wages Across Industries." *Review of Economics and Statistics*, 75(3), 409-417.

Neal, Derek A., and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy,* 104(5): 869-95.

Peski, Marcin, and Balazs Szentes. 2013. "Spontaneous Discrimination." *American Economic Review*, 103(6): 2412-36.

Reid, Alvin. 2015. "Black Head Coaches Don't Get Second Chances," *St. Louis American,* Aug 31, 2015, http://newpittsburghcourieronline.com/2015/08/31/black-head-coaches-dont-get-second-chances, accessed, January 24, 2016.

Rosen, Asa. 1997."An Equilibrium Search-Matching Model of Discrimination." *European Economic Review,* 41(8): 1589-1613.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50(1): 97-109.

Spencer, DeShuna. 2014. "For Black Youth in America, There Are No Second Chances," *emPower*, February 5, 2014, http://www.empowermagazine.com/black-youth-america-second-chances, accessed January 24, 2016.

Wozniak, Abigail. 2015. "Discrimination and the Effects of Drug Testing on Black Employment." *Review of Economics and Statistics*, 97(3): 548-566.

A Theoretical Appendix

A.1 Proof of Lemma 1

Proof. Define the quantities

ξ	Flow mass of workers born per unit time
A	Mass of unemployed black type α workers
B	Mass of unemployed black type β workers
Λ	Mass of currently monitored black type β workers

As g is the fraction of new workers that is type α and unemployed α workers are becoming employed each at a Poisson rate μ and never separate, A obeys

$$\frac{dA}{dt} = \xi g - \mu A$$

Similarly, a proportion $(1-g)$ of new workers is type β and such unemployed workers are also being hired at a Poisson rate μ each. However, as Λ workers who are of type β are being monitored, a flow mass $\Lambda\lambda(1-\beta)$ of black β workers are separating after monitoring reveals a bad match are also coming in to the black unemployed pool. Hence, B obeys

$$\frac{dB}{dt} = \xi(1-g) - \mu B + \Lambda\lambda(1-\beta).$$

Finally, unemployed β workers are becoming employed with monitoring at a Poisson rate μ and once they are employed they cease being monitored when match quality is revealed, which occurs at a rate λ. Thus the mass of monitored black β workers Λ must satisfy

$$\frac{d\Lambda}{dt} = \mu B - \Lambda\lambda.$$

Steady state implies that

$$\frac{dA}{dt} = \frac{dB}{dt} = \frac{d\Lambda}{dt} = 0.$$

Solving, we obtain

$$A = \frac{\xi g}{\mu}$$

$$B = \frac{\xi(1-g)}{\mu\beta}$$

and therefore the proportion of α workers in the unemployed pool is

$$\frac{A}{A+B} = \frac{\frac{\xi g}{\mu}}{\frac{\xi g}{\mu} + \frac{\xi(1-g)}{\mu\beta}} = \frac{g}{g + \frac{1}{\beta}(1-g)}.$$

Thus, a new match from the black job-seeker pool is of average quality

$$\frac{g}{g+\frac{1}{\beta}(1-g)} \cdot 1 + \left(1 - \frac{g}{g+\frac{1}{\beta}(1-g)}\right) \cdot \beta = \frac{\beta}{\beta g + (1-g)} \equiv \theta_B.$$

As $\beta < 1$ this is less than θ_W. ■

A.2 Proof of Lemma 2

Proof. Fix a wage w and a belief θ for the firm. If the firm doesn't monitor the worker, it receives the production net of the wage and the expected cost of errors, forever:

$$V^w_{\theta,N} = \frac{q - w - (1-\theta)\lambda c}{r}. \tag{15}$$

On the other hand, the firm can choose to monitor the worker. If it is optimal to do so at any instant, it is optimal to do so until the signal arrives as the problem doesn't otherwise change. With probability θ, the match is good and the production net of wages $q - w$ is received by the firm forever as no separation will occur; with the complementary probability the match is bad so $q - w - \lambda c$ is received by the firm only until the signal arrives and the match ends; in either case, the monitoring cost of b is paid until revelation. The firm's expected lifetime payoff if it monitors is therefore

$$V^w_{\theta,M} = \theta\frac{q-w}{r} + (1-\theta)\frac{q-w-\lambda c}{\lambda + r} - \frac{b}{\lambda + r}. \tag{16}$$

For monitoring to be optimal for the firm, we need

$$V^w_{\theta,M} \geq V^w_{\theta,N} \tag{17}$$

which reduces to

$$w \geq q - \lambda c + \frac{rb}{\lambda(1-\theta)}, \tag{18}$$

or equivalently

$$\theta \leq 1 - \frac{rb}{\lambda(w - q + \lambda c)}. \tag{19}$$

■

A.3 Proof of Lemma 3

Proof. As we require that strategies are not weakly dominated and form a PBE, the worker must at every wage draw be using an undominated action. As the worker's action doesn't affect his payoff if the firm rejects the wage, but does if the firm accepts it, she must always act as though the firm will accept the wage. Then, it follows that α workers will accept the wage offer if

$$w \geq rU_\theta^\alpha. \tag{20}$$

On the other hand, β workers will accept wages of

$$w \geq rU_\theta^\beta \tag{21}$$

regardless of whether they think the firm is likely to monitor and possibly fire them.[49] As α workers can mimic the β acceptance rule and not suffer separation, we have $U_\theta^\alpha \geq U_\theta^\beta$.

There are thus no wages α workers accept that β workers do not. If $U_\theta^\alpha > U_\theta^\beta$ there are wages that β workers accept that α workers do not; but then the firm will assign probability 1 to a worker accepting such a wage being type β and by $(C1)$ refuse such a wage. Thus no wage below rU_θ^α is ever accepted by both parties. On the other hand, since all workers accept higher wages, accepting such a wage does not shift beliefs; therefore, firms will accept all wages higher than rU_θ^α at which they make profits when they believe θ. The requirement that the firm makes nonnegative expected profit corresponds to at least one of $V_{\theta,M}^w$ and $V_{\theta,N}^w$ being positive; as they both are decreasing and continuous in w, we have that there is a single upper cutoff

$$\overline{w}_\theta = \max\{w | \max\{V_{\theta,M}^w, V_{\theta,N}^w\} \geq 0\}. \tag{22}$$

The fact that w_{θ_B} and w_{θ_W} are positive follows from C4 and the fact that both $V_{\theta,M}^w$ and $V_{\theta,N}^w$ are increasing in θ. Accordingly the lower cutoff is $\underline{w}_\theta = rU_\theta^\alpha$. ■

[49] Considerations about monitoring by other firms enter U_θ^β, but the worker's decision to accept a match does not depend on whether the firm will monitor.

A.4 Proof of Proposition 1

Proof. First, we show that in the white labor market, at the maximal wage $\overline{w}_{\theta_W}$, the employer does not monitor the worker. Suppose $w = q - (1-\theta_W)\lambda c$. Then we have

$$V^w_{\theta_W,M} = \theta_W \frac{q-w}{r} + (1-\theta_W)\frac{q-w-\lambda c}{\lambda + r} - \frac{b}{\lambda + r} \quad (23)$$

$$V^w_{\theta_W,M} = \frac{1}{r(\lambda + r)}\left[\theta_W(1-\theta_W)\lambda c - \frac{b}{\lambda}\right] \quad (24)$$

From C2 this expression is negative. As $V^w_{\theta_W,M}$ is decreasing in w, we must have that $\overline{w}_{\theta_W} = q - (1-\theta_W)\lambda c$. Therefore, no monitoring occurs at the upper end of the equilibrium wage interval. From Lemma 2 we have that the employer's monitoring decision is increasing in w; therefore, there is no monitoring in the market with belief θ_W. Given that, $\theta_W = \theta_0$ is the resulting steady state belief about newly hired workers.

The average wage in the market is therefore

$$w^{avg}_{\theta_W} = .5\overline{w}_{\theta_W} + .5\underline{w}_{\theta_W} = .5\overline{w}_{\theta_W} + .5\frac{\mu}{\mu + 2r}\overline{w}_{\theta_W} = \frac{\mu + r}{\mu + 2r}[q - (1-\theta_W)\lambda c]. \quad (25)$$

■

A.5 Proof of Proposition 2

Proof. Consider the wage at which the firm would break even if it does not monitor, $\overline{w}^n_{\theta_B}$. It is given by

$$q - (1-\theta_B) - \overline{w}^n_{\theta_B} = 0 \quad (26)$$

$$\overline{w}^n_{\theta_B} = q - (1-\theta_B)\lambda c. \quad (27)$$

Assume for contradiction that this is the highest equilibrium wage. As the monitoring decision is increasing in w, it must also not monitor at the lowest equilibrium wage, which by (3) is $\underline{w}^n_{\theta_B} = \frac{\mu}{\mu+2r}\overline{w}^n_{\theta_B}$. Furthermore, we have from (18) that such a non-monitoring wage must satisfy

$$\overline{w}^n_{\theta_B} \geq q - \lambda c + \frac{rb}{\lambda(1-\theta)}. \quad (28)$$

But that implies

$$\frac{b}{\lambda} \geq (1-\theta_B)\frac{\lambda c(\theta_B \mu + 2r) - 2rq}{r(\mu + 2r)}, \quad (29)$$

a direct contradiction to (C3). Therefore, we conclude that the firm would monitor at a wage of $\frac{\mu}{\mu+2r}\overline{w}^n_{\theta_B}$, that the highest equilibrium wage is the break-even monitoring wage $\overline{w}^m_{\theta_B}$, and that $\overline{w}^m_{\theta_B} > \overline{w}^n_{\theta_B}$. It then follows from the fact that $\frac{\mu}{\mu+2r}\overline{w}^n_{\theta_B}$ was a wage at which the firm would monitor and the increasing monitoring decision that the firm would monitor at $\frac{\mu}{\mu+2r}\overline{w}^m_{\theta_B}$, as well. Therefore, all workers are monitored and the average unemployed worker's quality is in steady state at θ_B.

The average wage in the back labor market is thus

$$w^{avg}_{\theta_B} = .5\overline{w}_{\theta_B} + .5\underline{w}_{\theta_B} = .5\overline{w}^m_{\theta_B} + .5\frac{\mu}{\mu+2r}\overline{w}^m_{\theta_B} = \frac{\mu+r}{\mu+2r}\left[q - \frac{r(\lambda c(1-\theta_B)+b)}{\lambda\theta_B + r}\right]. \quad (30)$$

■

B Empirical Appendix

B.1 Supervision

PSID Data We classify occupations into 12 groups of which nine are represented in the regression: professional (PSID occupation codes 10 through 19); not self employed (code 20); self employed (code 31, no one with this code is in the regression sample); clerks (codes 40 and 41); salesmen (code 45); craftsmen including foreman, nec (codes 50 and 51); protective services (codes 52 and 55, no one with these codes is in the regression sample); operatives (codes 61 and 62); laborers including farmers (codes 70, 71, and 80); service workers (codes 73 and 75); NA, Don't Know (code 99); and inapplicable (code 0, no one with this code is in the regression sample).

We classify industries into 11 groups, of which nine are represented in the regression: agriculture (PSID industry code 11); mining (code 21); construction (code 51); manufacturing (codes 30 through 49); transportation, communications, public utilities (codes 55 through 57); trade (codes 61 through 69); finance, insurance, real estate (code 71); services (codes 81 through 88); public administration (codes 91 and 92, no one with these codes in the regression sample); NA, Don't Know (code 99); inapplicable (code 0, no one with this code is in the regression sample).

We classify private employers based on the question regarding whether the respondent works for the federal, state, or local government. We classify those working for private employers as those who respond they do not work for the federal, state, or local government. We additionally code private employer as zero if the individual's reported industry was the armed services (code 91) or government, other than medical or educational services; NA whether other (code 92).

We classify individuals living in Alaska or Hawaii as living in the West region of the U.S.

Table B.1: Likelihood of Employer Monitoring by Race (coefficients on covariates)

Y = Level of Supervision	(1)	(2)	(3)	(4)
Black	0.150	0.212	0.176	0.185
	(0.099)	(0.104)	(0.110)	(0.110)
Other Race	-0.248	-0.225	-0.350	-0.334
	(0.176)	(0.199)	(0.208)	(0.211)
Highest Grade Completed			-0.063	-0.065
			(0.025)	(0.025)
Tenure (hundreds of months)			0.007	-0.012
			(0.155)	(0.158)
Tenure (hundreds of months)^2			-0.003	0.001
			(0.041)	(0.042)
Male			0.023	-0.017
			(0.128)	(0.135)
Age (tens of years)			-0.264	-0.310
			(0.205)	(0.207)
Age (tens of years)^2			0.016	0.021
			(0.024)	(0.024)
Union Job			0.037	0.004
			(0.109)	(0.110)
Salaried Worker			-0.390	-0.401
			(0.131)	(0.131)
Northeast			0.017	0.038
			(0.162)	(0.163)
North Central			-0.176	-0.159
			(0.147)	(0.146)
South			-0.127	-0.094
			(0.150)	(0.154)
Temporarily Laid Off			-0.072	-0.077
			(0.294)	(0.296)
Ln(Hourly Wage)				0.151
				(0.158)
Completed Education	≤ 12	≤ 12	≤ 12	≤ 12
Occupation, Industry FE	N	Y	Y	Y
N	1,095	1,095	1,089	1,089

Notes: Robust standard errors in parentheses. Estimates are from an ordered probit using data from the 1977 PSID. Dependent variable: level of employer supervision. A value of six corresponds to employer checking the individual's work several times per day, five to once a day, four to once a week, three to every few weeks, two to less often, and one corresponds to no supervisor. The sample includes household heads employed or temporarily laid off by private employers, who reported a wage and are not themselves supervisors. Observations are weighted by the family weights of the survey. See text for details.

B.2 Hazard Analysis Using the NLSY79

B.2.1 Constructing Individual Covariates

The stratified Cox estimation adjusts for highest grade completed at the time the employment spell begins. If the individual is surveyed in the year the employment spell begins, this is simply the highest grade completed that year. If the individual is not surveyed in that year, we impute the value of this variable as described below.

If the highest grade completed is the same in the previous and subsequent surveys, we impute the value for highest grade completed in the survey of interest. We first impute the values for survey years based on previous and subsequent surveys. After imputing these, we impute the value for nonsurvey years based on previous and subsequent surveys.[50]

Highest grade completed at the time the spell began will still be missing for individuals whose spell begins in nonsurvey years, and whose highest grade completed changes between the previous and the subsequent surveys.

If the individual is not surveyed in the year the spell begins, we determine the individual's geographic region at the beginning of the spell, and whether they live in an urban or rural location, using the value of these variables in surrounding years. We impute only if the value of these variables is the same in the previous and subsequent surveys.

B.2.2 Kernel- and Local-Linear Smoothing to Obtain Hazard Estimates: Methods

The principal results use intervals of weeks to smooth the hazard estimates. For robustness, we use week as a unit of time, the smallest unit of time for which we know employment status. We obtain the steps (hazard contributions) of the Nelson-Aalen cumulative hazard, and smooth them using a kernel smoother.[51] We obtain the cumulative hazard separately for black and white workers.

The cumulative hazard at time t_j is denoted $H(t_j)$. Then, the hazard contributions at each time t_j in which some individual is laid off are defined as:

$$\Delta\hat{H}(t_j) = \hat{H}(t_j) - \hat{H}(t_{j-1}) \tag{31}$$

[50] This addresses the problem arising from missing highest grade completed in 1998, but nonmissing in 1996 and 2000. Imputing the nonsurvey years and the survey years together, the value in 1997 would still be missing since the next survey year is 1998. Imputing the survey years first implies that 1998 becomes nonmissing, and so then 1997 becomes nonmissing as well because both 1996 and 1998 are nonmissing.

[51] The Nelson-Aalen cumulative hazard is $\hat{H}(t) = \sum_{j|t_j \leq t} \frac{d_j}{n_j}$, where d_j is the number of failures at time t_j and n_j is the number at risk of failure at time t_j. The steps of this function are equal to $\frac{d_j}{n_j}$.

We plot the smoothed hazard function separately for black and white workers

$$\hat{h}(t) = b^{-1} \sum_{j=1}^{D} K_t(\frac{t - t_j}{b}) \Delta \hat{H}(t_j) \tag{32}$$

The hazard estimate at time t, $h(t)$, is based on the the hazard contributions $\hat{H}(t_j)$ at all failure times j, where each contribution is weighted by the kernel function K and bandwidth b. We use the Epanechnikov kernel, with the bandwidth equal to .5*Silverman's plug-in estimate.[52] This yields a bandwidth of 67 for black workers and 72 for white workers.

We use a boundary-adjusted Epanechnikov kernel (based on Müller and Wang (1994)) to address bias in the boundary regions ($t_{min} \leq t < b$; $t_{max} - b < t \leq t_{max}$) from using a symmetric kernel. As a further alternative to using a boundary-adjusted kernel, we show hazard estimates only outside the boundary region. For those results, we smooth the hazard contributions using the version of the Epanechnikov kernel described in Epanechnikov (1969).[53] This yields equivalent results to the version of the Epanechnikov kernel described in the previous paragraph and in Footnote 52 if the bandwidth is divided by $\sqrt{5}$. This implies the boundary region is smaller when using this kernel, which is helpful given we plot only outside the boundary. These plots show a clear gap at the boundary of roughly 30 weeks which closes over time.

We also smooth the hazards using kernel-weighted local linear regression which does not yield biased estimates in the boundary regions (see Nielsen and Tanggaard 2001 and Cameron and Trivedi 2005).

We additionally estimate Cox models stratified by race, and use the same Epanechnikov kernel and bandwidth as in the nonparametric specifications to smooth the baseline hazard contributions from the Cox model. These models include the same covariates as the models using week bins.

B.2.3 Coding Quits

The respondents' choices for why they left their job change with the surveys. For most years the reasons why the respondent left their job include several which are explicitly named

[52] The version of the Epanechnikov kernel we refer to here is $K[z] = .75(1 - z^2)$ if $|z| < 1$. Silverman's plug-in estimate is given by $b^* = 1.3643\delta N^{-.2}\min(s, iqr/1.349)$ where δ depends on the kernel and is 1.7188 for the Epanechnikov kernel, N is the number of unique failure times, s is the sample standard deviation of the failure times and iqr is the interquartile range of the failure times. Cameron and Trivedi (2005) suggest using Silverman's plug-in estimate, as well as bandwidths half and twice the size. Because Silverman's plug-in estimate for the bandwidth is quite large, we present estimates using bandwidths half the size as well as the actual Silverman plug-in estimate.

[53] $K[z] = .75(1 - \frac{1}{5}z^2)$ if $|z| < \sqrt{5}$

"quits" for example quit for pregnancy or family reasons or quit to look for another job, and several which are clearly not quit related. In the first survey year, as well as the surveys starting in 2002, there are several additional categories that do not include the word "quit" that we code as quits as they are arguably voluntary separations.

In 1979, these include "Pregnancy" and family changed jobs or moved, or family reasons. In years after 2002, these include "went to jail, prison, had legal problems," "transportation problems," "no desirable assignments available," "retirement," "job assigned through a temp agency or a contract firm became permanent," and "dissatisfied with job matching service". There are few instances of spells ending for these reasons.

B.2.4 Occupation and Industry

We use data on occupation from the Employer Roster. Up until 2000, the 1970 Census occupation codes are used in the NLSY to code occupation. For only some jobs, the 1980 Census occupation codes are used starting in 1982. Starting in 2002, the 2000 Census occupation codes are used. These are used through the 2004 survey, with the exception that in the 2004 NLSY survey there is an added zero on each of the occupation codes, which matches the added zero in the Census ACS codes starting in 2003. Starting in 2006, the ACS 2005-2009 occupation codes are used in the NLSY to code occupation.

In order to obtain consistent occupation codes across surveys, we use the crosswalks from the Minnesota Population Center to convert the occupation codes to 1990 Census Occupation Codes. Specifically, we download the 1970 Census state, metro, and neighborhood samples, with both the contemporaneous Occupation Codes and the 1990 Occupation codes. Keeping one observation per contemporaneous occupation code, we merge on contemporaneous occupation code with the NLSY occupation roster, only to observations in survey years before 2002 since those were based on 1970 codes.

We then download the 2000 Census 1 percent sample with both the contemporaneous occupation codes and the 1990 occupation codes, and implement the same procedure for survey years 2002 and 2004. We download the 2005 ACS with the contemporaneous occupation codes and the 1990 occupation codes, and implement the same procedure for years starting in 2006.

Finally, we download the 1980 census 5 percent state sample, and 1 percent metro, urban, LMA, and metro/non-metro samples. For observations from surveys before 2002 that did not merge with the 1970 census, we then merge with the 1980 census in case any of these nonmerged occupations were 1980 codes rather than 1970 codes.

We then classify occupations into 19 groups, which correspond to the subclassifications of the seven occupation groups suggested by IPUMS. These groups include: executive, administrative, and managerial occupations; management related occupations; professional specialty occupations; technicians and related support occupations; sales occupations; administrative support occupations, including clerical; private household occupations; protective service occupations; service occupations, except protective and household; farm operators and managers; other agricultural and related occupations; mechanics and repairers; construction trades; extractive occupations; precision production occupations; machine operators, assemblers, and inspectors; transportation and material moving occupations; military occupations; unemployed and missing. Due to the small number of individuals in military occupations given our sample restrictions, we group these occupations with unemployed and missing, consistent with IPUMS grouping these together as "non-occupational responses". This results in 18 occupation groups.

Individuals with missing occupation include those whose occupational code in the NLSY did not match to a 1990 Census occupation code, those whose NLSY occupation codes match to a 1990 Census occupation code indicating N/A and unknown or unemployed, and those who were not in the occupation roster data. Among those whose occupational code in the NLSY did not match to a 1990 Census occupation code, nearly all of these are because the occupation code in the NLSY was -4, the code for a valid skip. There were several individuals whose NLSY occupational code was outside of the non-occupational and invalid responses, but did not match a 1990 Census occupation code. For these individuals there was no contemporaneous occupation code matching the NLSY occupation code.

The Minnesota Population Center proposes the following more aggregated set of occupational categories: managerial and professional; technical, sales, and administrative; service; farming, forestry, and fishing; precision production, craft, and repairers; operatives and laborers; and non-occupational responses (Minnesota Population Center 2020).

We follow an analogous procedure for obtaining the industry codes. Up until 2000, the 1970 Census industry codes are used in the NLSY to code industry. For only some jobs, the 1980 Census industry codes are used starting in 1982. Starting in 2002, the 2000 Census industry codes are used. Starting in 2004, the 2002 Census industry codes are used. To obtain the crosswalk between the 1990 Census industry codes and the 2002 Census industry codes, we download the 2003 through the 2007 ACS with both the contemporaneous and 1990 industry codes. These are the samples in which the contemporaneous industry code is the 2002 Census industry code.

We classify industries into 15 groups, corresponding to those delineated by the Minnesota Population Center. These include: agriculture, forestry, and fisheries; mining; construction; manufacturing; transportation, communication, and other public utilities; wholesale trade; retail trade; finance, insurance, and real estate; business and repair services; personal services; entertainment and recreation services; professional and related services; public administration; military; and missing. Based on our sample restrictions, there are not any individuals reporting the military as their industry. This results in 14 industry groups.

Individuals with missing industry code include those whose industry code in the NLSY did not match to a 1990 Census industry code, those whose industry code in the NLSY matches to a 1990 Census industry code indicating N/A, did not respond, or last worked in 1984 or earlier, and those who were not in the industry roster data.

Among those whose industry code in the NLSY did not match to a 1990 Census industry code, nearly all of these are because the industry code in the NLSY was -4, the code for a valid skip. There were a small number of individuals whose NLSY industry code was outside of the non-industry and invalid responses, but did not match a 1990 Census industry code. For these individuals there was no contemporaneous industry code matching the NLSY industry code.

References

Cameron, A. Colin and Pravin K. Trivedi. 2009. *Microeconometrics: Methods and Applications*, 8th edition. Cambridge, England: Cambridge University Press.

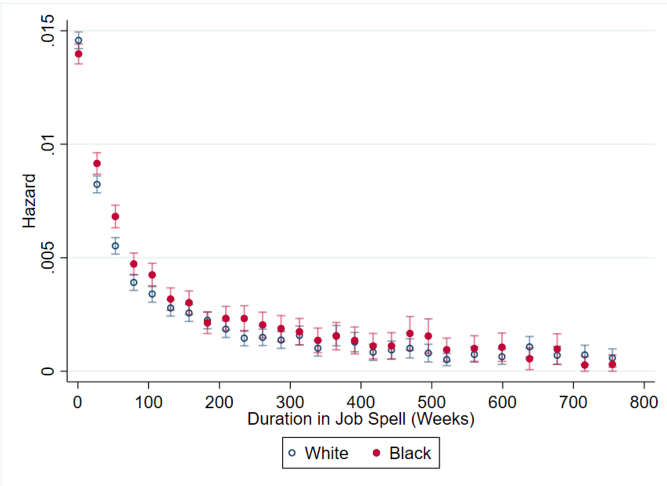
Epanechnikov, V. A. 1969. "Non-Parametric Estimation of a Multivariate Probability Density." *Theory of Probability and its Applications,* 14(1): 153-158.

Minnesota Population Center. "OCC1990: Description," https://usa.ipums.org/usa-action/variables/OC access March 23, 2020.

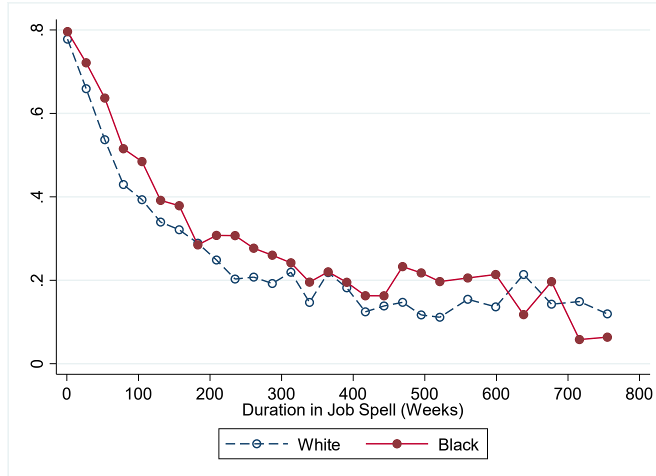
Nielsen, Jens Perch and Carsten Tanggaard. 2001. "Boundary and Bias Correction in Kernel Hazard Estimation." *Scandinavian Journal of Statistics*, 28: 675-698.

Appendix Figure B1: Robustness Hazard Estimates using Week Bins

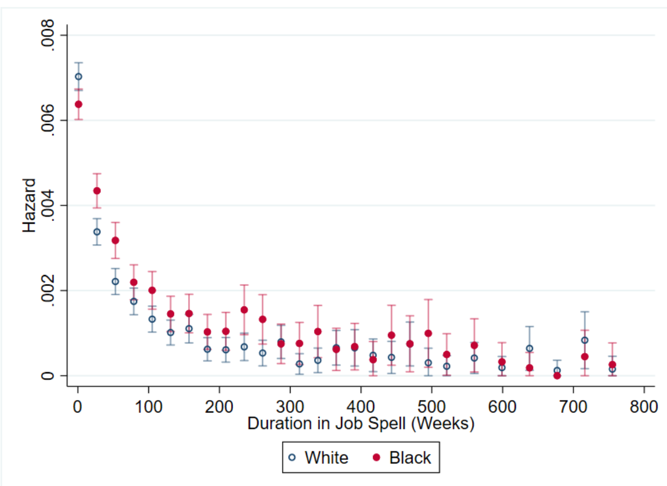
(a) Treating Quits into Nonemployment Similarly to Layoffs, Nonparametric



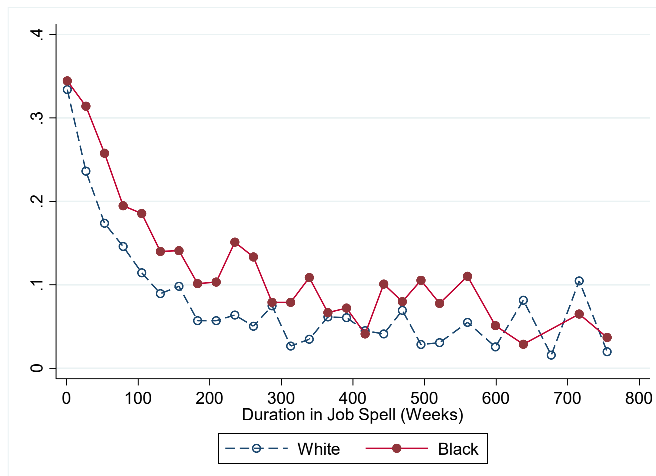
(b) Treating Quits into Nonemployment Similarly to Layoffs, Cox



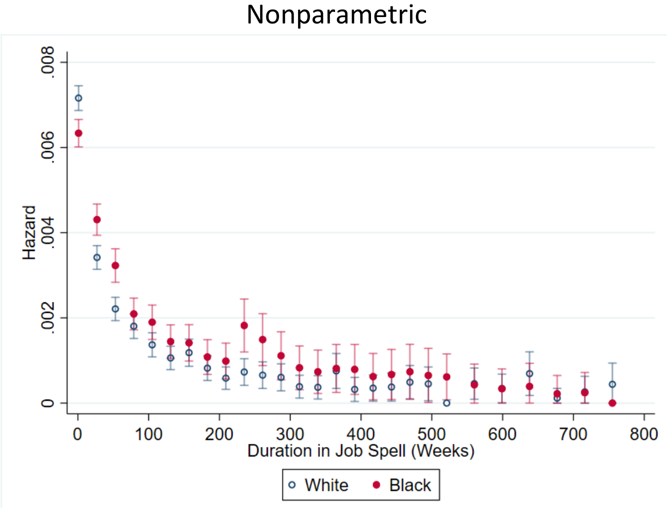
(c) Enter Spell from Nonemployment, Nonparametric



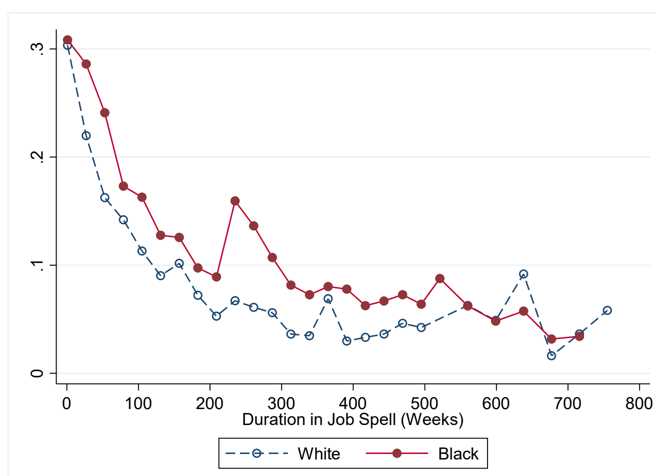
(d) Enter Spell from Nonemployment, Cox



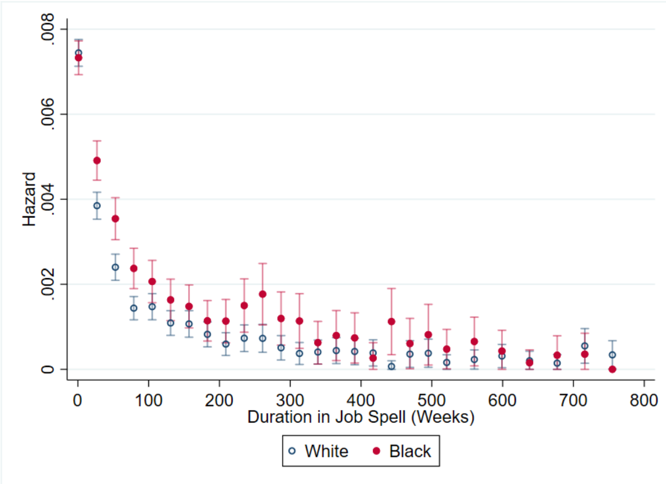
(e) Any Interruption at Employer as Ending Spell, Nonparametric



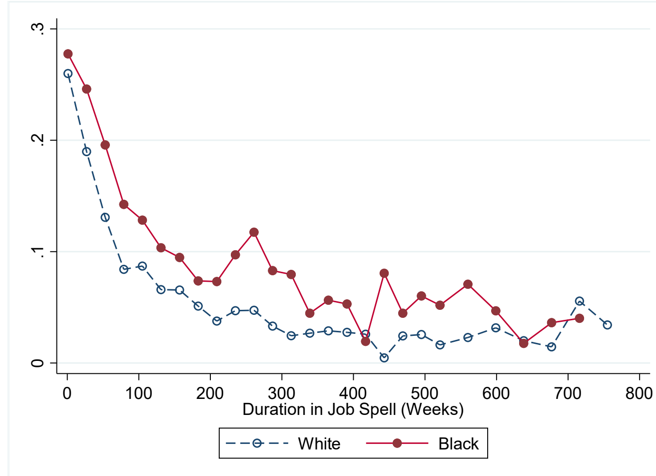
(f) Any Interruption at Employer as Ending Spell, Cox



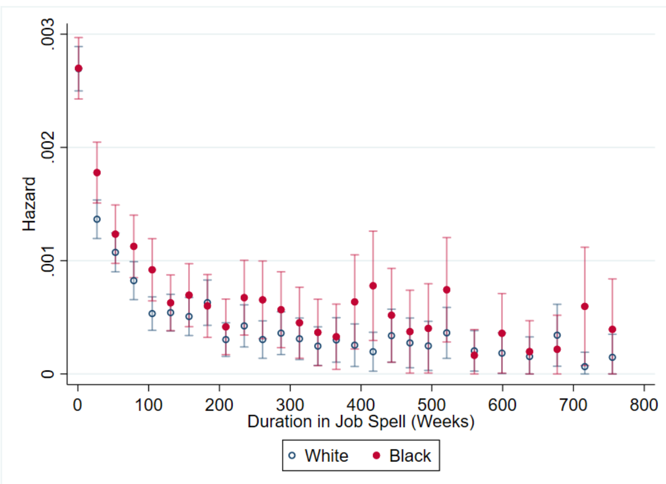
(g) Age ≤ 30 at Spell Start, Nonparametric



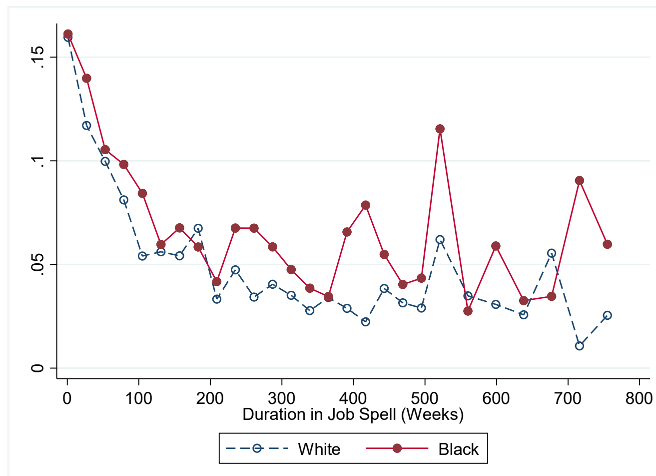
(h) Age ≤ 30 at Spell Start, Cox



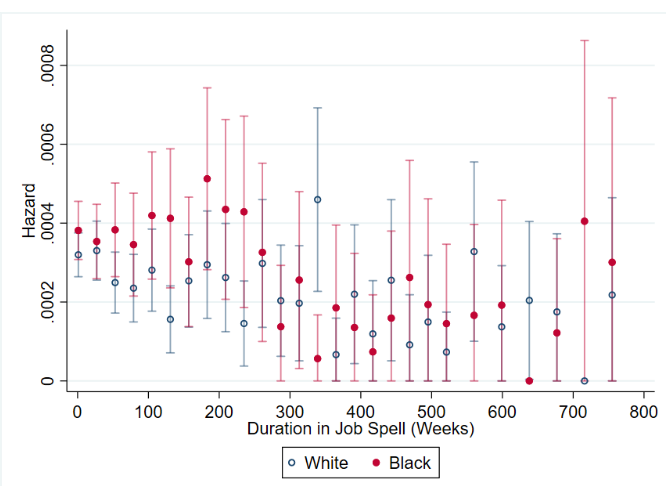
(i) > 12 Years of Education at Spell Start, Nonparametric



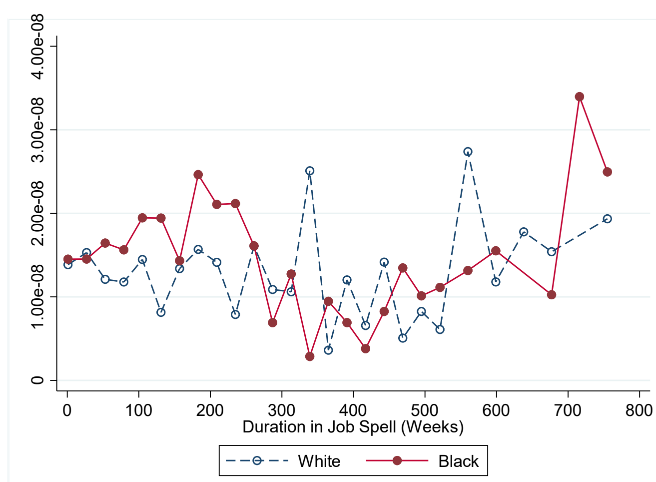
(j) > 12 Years of Education at Spell Start, Cox



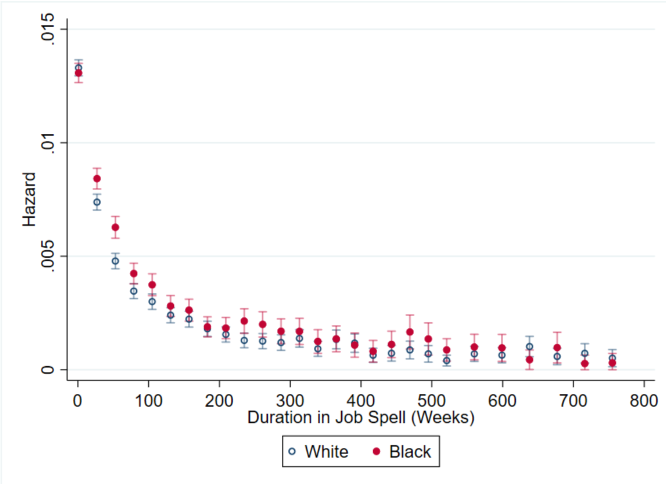
(k) Hazard of Job Ending due to Plant Closing, Nonparametric



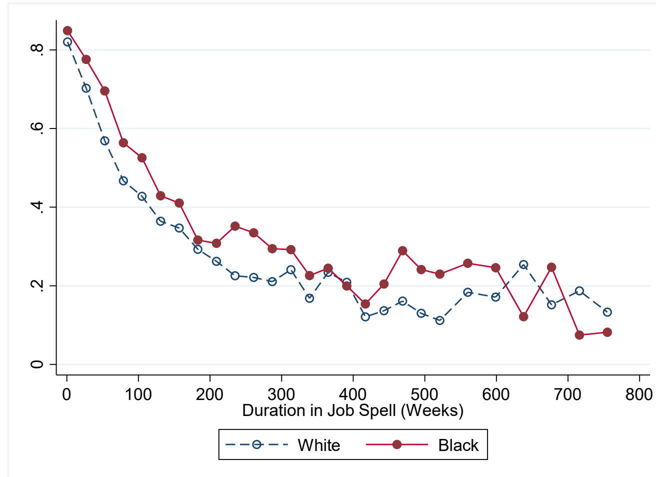
(l) Hazard of Job Ending due to Plant Closing, Cox



(m) Quits to Nonemployment as Layoffs, End in Nonemployment if > 4 Weeks Until Next Fulltime Spell, Nonparametric



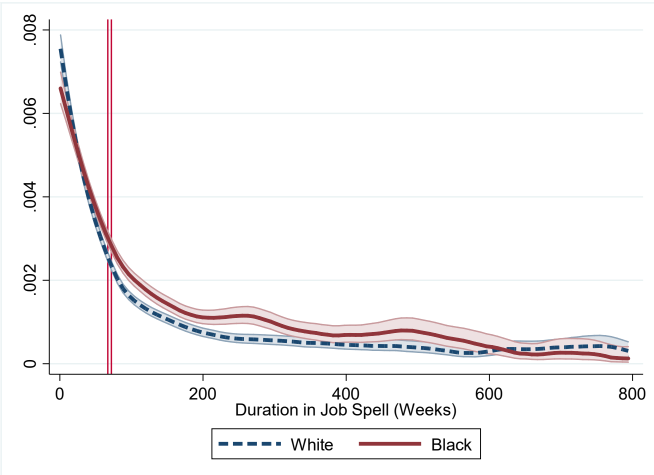
(n) Quits to Nonemployment as Layoffs, End in Nonemployment if > 4 Weeks Until Next Fulltime Spell, Cox



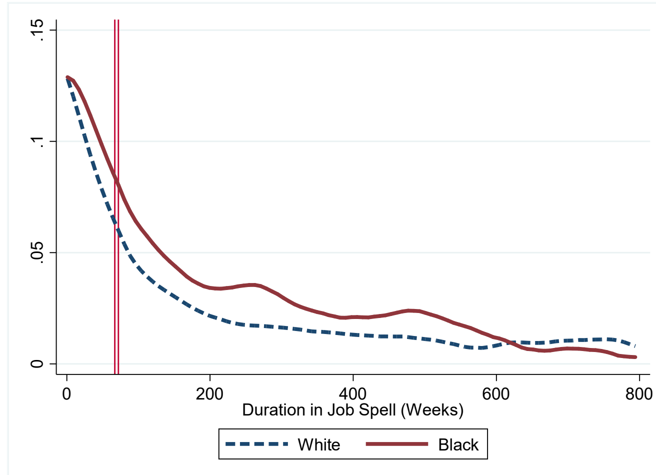
Notes: This figure presents robustness results, using both nonparametric methods and stratified Cox regressions. These are the same methods as used in the principal results described in Figures 4 and 5. The plots in (a) and (b) treat quits into nonemployment similarly to layoffs, where nonemployment is defined as more than one week between fulltime spells. The plots in (c) and (d) restrict to employment spells for which the individual entered from nonemployment, which includes an individual's first fulltime spell. The plots in (e) and (f) treat any gap at the same employer as ending the spell, rather than treating gaps of ≤ 26 weeks at the same employer as the same spell. The plots in (g) and (h) restrict to spells for which the individual was less than or equal to 30 at the start of the spell. The plots in (i) and (j) restrict to individuals who completed more than 12 years of education at the start of the spell. The plots in (k) and (l) present the hazard of a job ending due to a plant closing. The plots in (m) and (n) treat quits into nonemployment similarly to layoffs, where nonemployment is defined as more than four weeks between fulltime spells. For all plots showing confidence intervals, bands around estimates are 95% confidence intervals.

Appendix Figure B2: Robustness Kernel-Smoothed Hazard Estimates

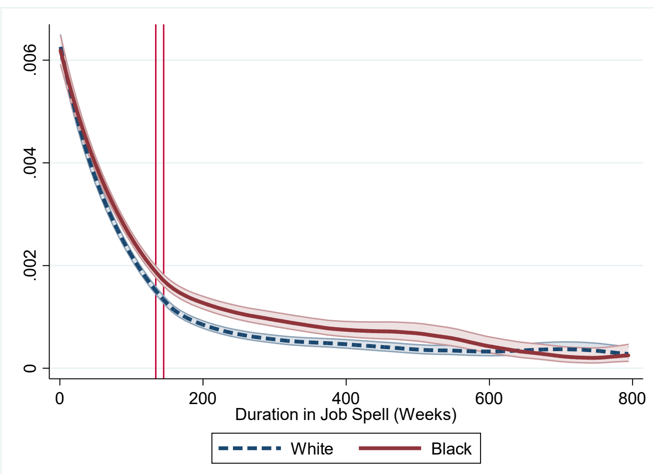
(a) Boundary-adjusted, Smaller Bandwidth, Nonparametric



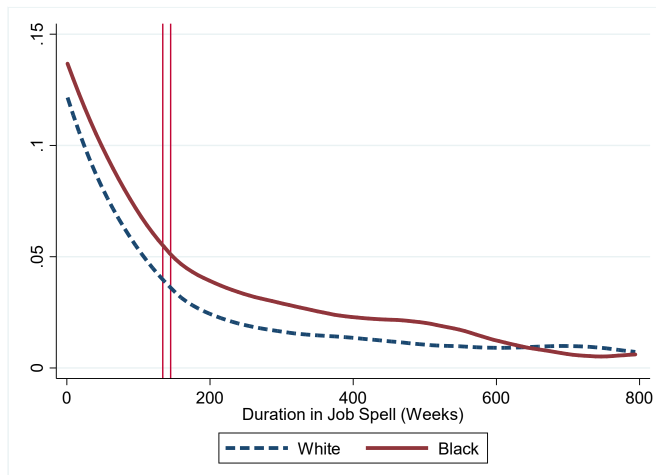
(b) Boundary-adjusted, Smaller Bandwidth, Cox



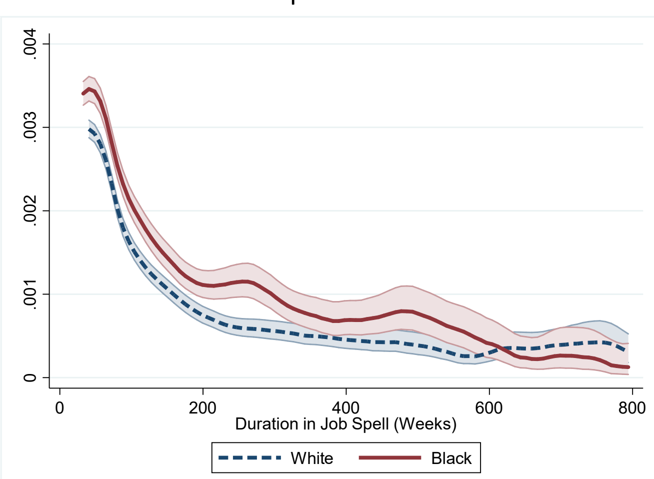
(c) Boundary-adjusted, Larger Bandwidth, Nonparametric



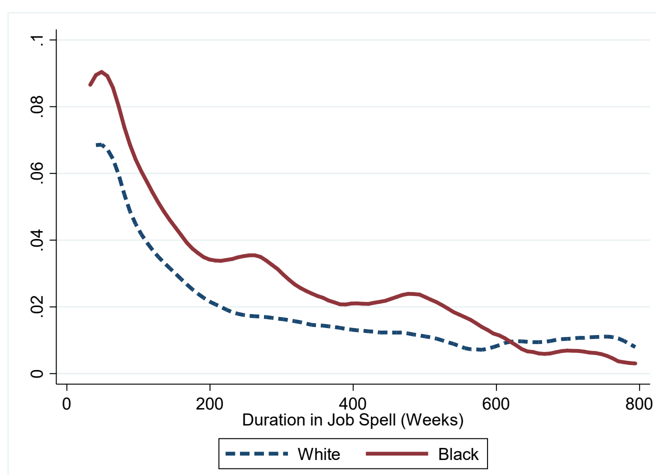
(d) Boundary-adjusted, Larger Bandwidth, Cox



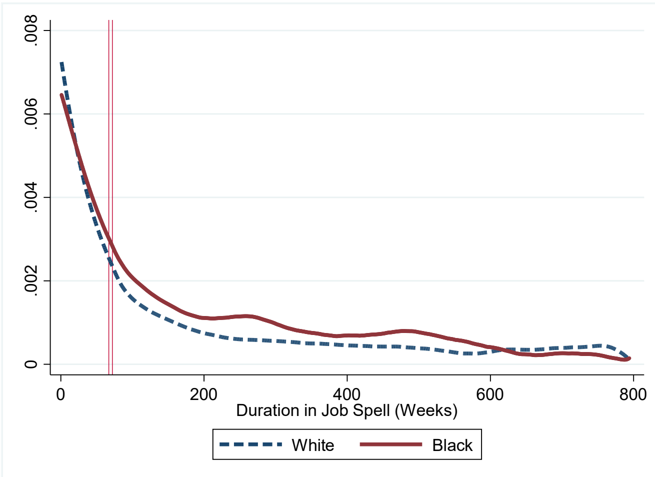
(e) Exclude Boundary Region, Smaller Bandwidth, Nonparametric



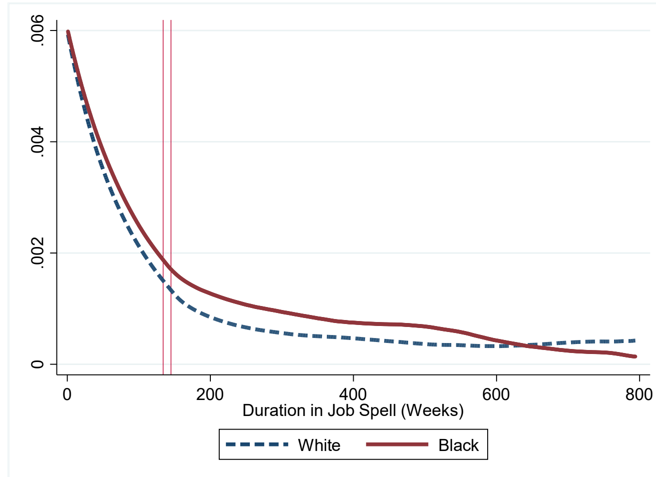
(f) Exclude Boundary Region, Smaller Bandwidth, Cox



(g) Local Linear Smoothing, Smaller Bandwidth, Nonparametric

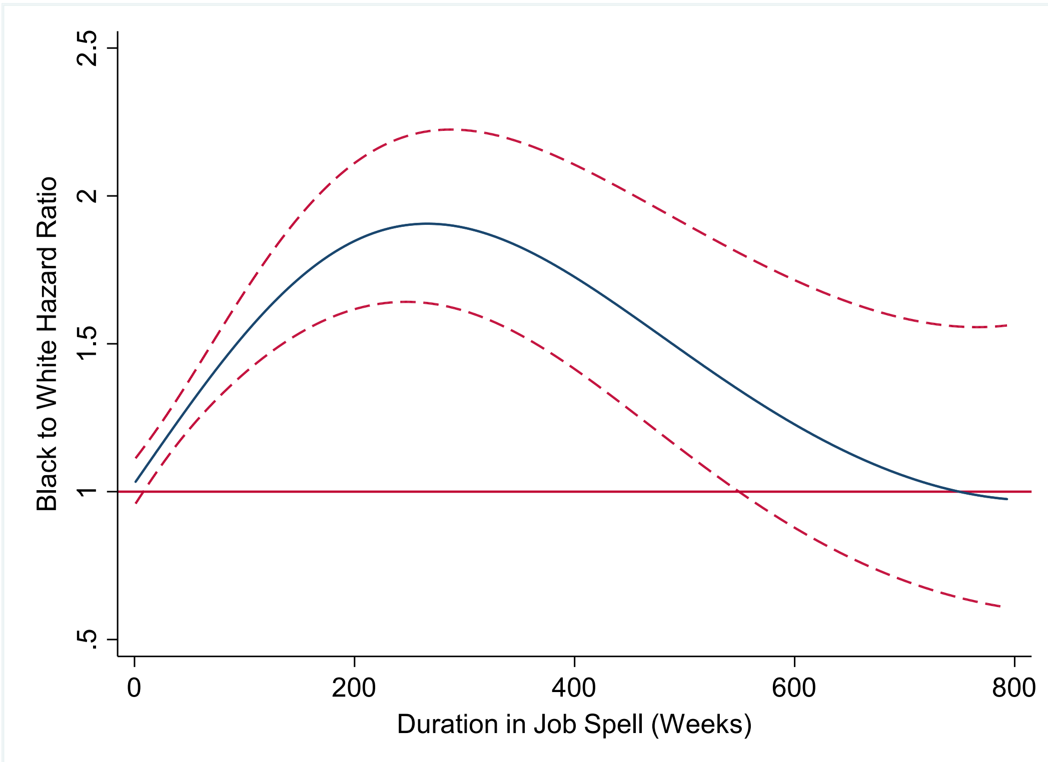


(h) Local Linear Smoothing, Larger Bandwidth, Nonparametric



Notes: The plot in (a) presents kernel-smoothed hazard contributions from the Nelson-Aalen cumulative hazard, using the boundary-adjusted alternative Epanechnikov kernel, with bandwidth equal to 72 weeks for white workers and 67 weeks for black workers, these are one half of the Silverman plug in for this kernel. The plot in (b) presents kernel-smoothed hazard contributions from a Cox Proportional Hazards model, stratified by race. The kernel and the bandwidth are the same as those in (a). The explanatory variables in the Cox model are the same as those included in Figure 5. The plots in (c) and (d) are analogous to (a) and (b), but use the Silverman plug-in bandwidth (145 and 134 weeks for white and black workers respectively) rather than half of the Silverman plug-in bandwidth as in (a) and (b). The plots in (e) and (f) are analogous to (a) and (b), but use the Epanechnikov kernel (1969 version), with bandwidth equal to half the Silverman plug in for this kernel (bandwidth of 32 weeks for white workers and 30 weeks for black workers), and show results only outside the boundary regions. The plot in (g) uses local linear smoothing of the hazard contributions, using the alternative Epanechnikov kernel with bandwidth equal to half the Silverman plug-in estimate. The plot in (h) is the same as in (g), but uses the Silverman plug-in estimate as the bandwidth. For all plots showing confidence intervals, bands around estimates are 95% confidence intervals. Vertical lines show boundary regions for black and white workers.

Appendix Figure B3: Hazard Ratio for Black Workers Relative to White Workers Controlling for Covariates in a Cox Model, Allowing the Percentage Gap in Hazards to be a Cubic in Seniority



Notes: This is a plot of the hazard ratio for black relative to white workers from a Cox model controlling for the same covariates included in Figure 5. Additionally, we include an indicator for whether the worker is black, and interact this with a cubic in seniority (duration in job spell in weeks). In order to allow the effect of race to vary over time, we include an observation for each job spell at each failure time in the data (as Cox models are only estimated when failures occur in the data). We obtain the linear combination of the coefficients on the indicator for black worker, for values of week from 1 to 793. We then exponentiate these to obtain the hazard ratio. Dashed lines are the 95% confidence intervals based on robust standard errors, which are larger than the nonrobust standard errors.

Appendix Table B2: Coefficients from Cox Model Stratified by Race

Highest Grade Completed at Spell Start	-0.0668
	(0.0102)
Male	0.182
	(0.0292)
AFQT (Percentile)	-0.00208
	(0.000587)
AFQT (Percentile) Missing	-0.215
	(0.0700)
Region at spell start: Northeast	-0.233
	(0.0424)
Region at spell start: North Central	-0.143
	(0.0380)
Region at spell start: South	-0.297
	(0.0373)
Region at spell start: Missing	-0.180
	(0.133)
Urban Location at Spell Start	0.0125
	(0.0295)
Urban Location at Spell Start Missing	-0.0213
	(0.0734)
Age at Spell Start	-0.0151
	(0.00580)
Observations	33,814

Notes: Conventional standard errors in parentheses; these are larger than the robust standard errors which account for observations appearing multiple times in risk pools. Coefficients are from a Cox Proportional Hazards model stratified by race, using week bin as a unit of time. Each observation is a job spell. We model the layoff hazard, and the failure variable is an indicator for whether the spell ended because the individual was fired or laid off. The regression additionally includes fixed effects for year, occupation (18 groups), and industry (14 groups) all measured at the start of the spell. See text for details.

Appendix Table B3: Differential Effect of AFQT on the Layoff Hazard, by Race

	White	Black	All
AFQT (Percentile)	-0.0012	-0.0049	-0.0012
	(0.0007)	(0.0012)	(0.0007)
AFQT*Black			-0.004
			(0.0013)
Black			-0.188
			(.222)
Highest Grade Completed at Spell Start	-0.078	-0.051	-0.082
	(0.014)	(0.016)	(0.014)
Highest Grade Completed at Spell Start*Black			0.043
			(0.020)
Male	0.236	0.108	0.182
	(0.039)	(0.045)	(0.030)
Region at spell start: Northeast	-0.24	-0.100	-0.233
	(0.051)	(0.087)	(0.043)
Region at spell start: North Central	-0.165	-0.004	-0.148
	(0.043)	(0.087)	(0.039)
Region at spell start: South	-0.38	-0.107	-0.306
	(0.045)	(0.078)	(0.038)
Urban Location at Spell Start	0.0200	0.013	0.017
	(0.037)	(0.050)	(0.030)
Age at Spell Start	-0.022	-0.004	-0.017
	(0.008)	(0.009)	(0.006)
Observations	20,140	13,674	32,627

Notes: Conventional standard errors in parentheses. Coefficients are from a Cox Proportional Hazards model, using week bin as a unit of time. Each observation is a job spell. We model the layoff hazard, and the failure variable is an indicator for whether the job spell ends because the individual was fired or laid off. The regression additionally includes fixed effects for year, occupation (18 groups), and industry (14 groups) all measured at the start of the spell, as well as indicators for whether AFQT, region, and urban location are missing. Column 3 excludes individuals with missing AFQT. See text for details.