Beliefs and Affirmative Action in Employment*

Benjamin S. Griffy†
University at Albany, SUNY

Eric R. Young‡
University of Virginia

Federal Reserve Bank of Cleveland

This version: November, 2022

Abstract

In this paper, we show that when firms and workers both face asymmetric information, Black workers may attribute poor employment outcomes to taste-based discrimination that actually results from statistical discrimination by firms. This leads to less investment in human capital by Black workers, who expect that scant employment prospects will yield little return, which simultaneously reinforces the firms belief about the qualifications of Black workers. We show that in this circumstance, Affirmative Action can increase Black human capital investment by revealing information about the degree of discrimination they face, which can cause long-term increases in human capital investment even if the Affirmative Action policy is enacted only for a short period of time.

JEL Classification: E21, E24, J63, J64, D31, I32, J31
Keywords: Affirmative Action, Inequality, Race, Racial Inequality, Macroeconomics, Information Frictions

*We thank seminar participants at the New York Federal Reserve, the MacCalm Conference hosted by the University of Edinburgh, the Bureau of Labor Statistics, the Penn Graduate Student Workshop, and the COCONUTS Virtual Search Seminar.

†Correspondence: Department of Economics, SUNY Albany, Room BA110, 1400 Washington Ave., Albany, NY, 12222. Tel.:(503)-341-3950. Email: bgriffy@albany.edu. Web: https://www.bengriffy.com

‡Correspondence: Department of Economics, University of Virginia, 248 McCormick Road, Charlottesville VA 22911. Tel: (434) 924-3811. Email: ey2d@virginia.edu. Web: https://sites.google.com/view/ericyoung/home

1 Introduction

Black Americans lag behind their White peers in nearly every measure of economic achievement: they earn less on average, endure more frequent unemployment spells, are underrepresented in prestigious occupations, and within these prestigious occupations, they earn less. These facts raise an obvious question: can Affirmative Action policies increase participation in better occupations, and can they create sustained improvements? Affirmative Action was once a widely-adopted suite of policies intended to fight taste-based discrimination by employers, but intellectual headwinds changed and political will ebbed. The case against Affirmative Action centers on the notion that competition would solve to economic ills targeted by these policies: over time, discriminating employers would be competed out of the market by reducing their pool of workers and consumers (Becker, 1957). Even worse, hiring underqualified Black workers as mandated by these policies could change employers' beliefs so that they are less likely to hire Black workers in the future (Arcidiacono et al., 2011). This argument, however, ignores the positive effects that Affirmative Action can have on future generations of Black workers by revealing that they can succeed in careers that were they previously deemed unobtainable.

In this paper we address a crucial question about this beliefs channel: can temporary Affirmative Action policies have long-lasting positive effects for Black Americans by changing beliefs about employment prospects and inducing human capital investment? We find that, indeed, Affirmative Action even over a short horizon permanently reshapes the beliefs of future generations and causes sizable gains in skill investment. This occurs even when Affirmative Action initially displaces more qualified White workers and causes firms to reduce their beliefs about the qualifications of Black candidates. In our model, workers and firms face information asymmetries: Black workers believe a large share of firms are taste-based discriminators and all but the most capable choose to forgo human capital, believing that this investment would prove fruitless. As a consequence, even non-discriminatory firms

considering Black applicants place little faith in positive, but noisy, information garnered from the interview unless the candidate is exceptional. The resulting absence of Black hires leaves firms with little reason or opportunity to revise their beliefs, and reinforces Black workers' beliefs about the extent of discriminatory employers. In this context, an Affirmative Action policy proves vital to address these information problems: by requiring firms to place additional weight on an applicants' race, non-discriminatory firms hire more Black workers, which changes the beliefs of the next generation of Black workers. While firms revise down their beliefs about the skills of future Black workers, the next generation of Black workers begin to believe that investment in human capital will yield a return because by observing the success of the affected generation, they learn that there are fewer taste-based discriminators than believed and that the remaining firms are willing to hire Black workers who appear qualified. This results in more investment by future generations of Black workers, which becomes self-reinforcing as firms learn that the average Black applicant is more capable than they previously believed. With the addition of this beliefs channel Affirmative Action proves crucial: in the absence of such policies, racial equity may languish or improve at a pace so glacial that it may never be realized.

We first review evidence about racial discrimination against Black workers and document the impact it has on their beliefs. We then use this evidence as the backdrop for the construction of our model. We find evidence that the beliefs play a crucial role in career aspirations: Black youths surveyed in the National Longitudinal Survey of Youth 1979 (NLSY79) aspire to both less prestigious and lower pay careers when they believe discrimination will affect their labor market in the future. Consistent with our model, the less-prestigious occupations to which Black youths who fear discrimination aspire exhibit higher concentrations of Black workers at the time of the survey. We also document a widespread belief among Black workers in the 1970s that discrimination negatively affected their labor market as well as evidence of racial animus among supervisors with hiring discretion using the General Social Survey (GSS).

We use this evidence to inform our structural model. We construct a model in which both firms and workers face asymmetric information. Workers live for one period and may be Black or White. They draw from an ability distribution and make a subsequent human capital investment decision based on their innate ability and their beliefs about their employment prospects, which may differ by race. Both races hold a common belief about taste-based discriminatory firms in the economy, which influences their beliefs about their employment probability for a given level of human capital investment. A new cohort is born each period and updates this belief after they observe the employment outcomes of the previous cohort. Some firms are taste-based discriminators and all firms start each period with beliefs about the distribution of human capital by race. They use these beliefs to make a hiring decision for a pool of Black and White applicants, whose human capital is only fully revealed after being hired.

Using our model, we assess the impact of Affirmative Action on employment outcomes and human capital investment over time. We find that installing Affirmative Action policies for a single generation closes the income gap between White and Black workers by 14 percentage points (56.6%) and reduces the employment gap in high prestige occupations by 39 percentage points (67%). This occurs because the generation after the period of Affirmative Action reduces their belief about the measure of discriminatory firms (68 percentage points), and as a consequence, this cohort of Black workers increases their human capital investment by 41.5%. This persists as firms interview and hire new Black candidates, causing them to adjust their beliefs to reflect this new investment. These results occur despite our finding that imposing Affirmative Action can negatively affect employers' perceptions about potential Black employees: 58% of firms hire a less qualified Black applicant than the White applicant they hired in the Baseline, while only 2.6% hire a more qualified Black applicant. However, this effect is transitory and dissipates within a period as subsequent cohorts increase their investment.

Next, we demonstrate the importance of allowing for two-sided beliefs that update when

analyzing the impact of Affirmative Action. We first consider an environment in which only firm beliefs are able to update. Because worker beliefs affect their investment decisions, Affirmative Action yields negligible benefits in our environment when only firms may update their beliefs. Here, workers never respond to the revelation of a favorable labor market, leading to only minimal revisions by firms. We then consider an environment in which only worker beliefs are able to update. While we find that in this environment, Affirmative Action can still positively affect outcomes, statistical discrimination becomes permanently embedded in Black workers beliefs, leading to an economy with persistently less investment and a belief that the labor market is less favorable for Black workers.

Last, we show that Affirmative Action is not a universal panacea for racial inequality: in certain belief settings, imposing Affirmative Action policies can cause long-term harm to Black workers. In our analysis, circumstances in which initial discrimination beliefs do not adequately account for the presence of statistical discrimination lead to negative consequences of imposing Affirmative Action policies. This is because the information revealed by Affirmative Action indicates that previous beliefs about discrimination were understated and the labor market is more hostile than previously believed. We show examples of two settings in which this may occur: one in which the measure of taste-based discriminators is large, and one in which initial beliefs about discrimination reflect only the measure of taste-based discriminators. In both examples, Affirmative Action leads to persistently worse outcomes for Black workers, because it reveals that their beliefs understated the degree of discrimination in the labor market.

While the channel that we introduce in this paper is perhaps not novel to casual discourse, it is novel to the literature on statistical discrimination and Affirmative Action. Previous work typically considered static environments (see Coate and Loury (1993), Moro and Norman (2004), and Fang and Norman (2006) among others) or environments in which Affirmative Action dispels statistical discrimination (see Lundberg and Startz (1983)).[1] Among the

[1]See Fang and Moro (2011) for an exhaustive review of the theories of statistical discrimination and

limited number of papers that consider a similar dynamic environment are Blume (2006) and Levin et al. (2009). Both papers allow for firms and workers to update beliefs in response to labor market outcomes. In both papers, however, discrimination is a purely statistical process that results from mismatch in the labor market and is common knowledge to firms and workers. In this context, Affirmative Action would have no role because workers know the measure of taste-based discriminators and are certain about all the primitives of their environment. Another related paper, Chung (2000), assesses the effect of Affirmative Action on human capital investment through increasing representation. In his paper, however, Affirmative Action necessarily causes a switch in beliefs because the hiring decision is not modeled and workers update their beliefs only on aggregate outcomes; here, we model the hiring decision and show circumstances in which Affirmative Action can be a negative. Among more recent papers, Boerma and Karabarbounis (2021) consider the effect of reparations on wealth inequality using a model in which beliefs about the return to capital investment depends on past experiences. They find that reparations alone cannot close the gap due to beliefs, which complements our findings about the importance of beliefs.

Our findings also complement the results from the literature studying the impact of Affirmative Action in college admissions. This literature is divided on the effects of Affirmative Action (see Fischer and Massey (2007) for an example casting Affirmative Action in a positive light and Sander (2004) for the opposite, among many others), but these papers rarely consider the dynamic effects on subsequent cohorts, which we highlight as the most important feature of Affirmative Action in our paper. Among the most pertinent for our analysis is Bagde et al. (2016) which considers the long-term effects of Affirmative Action programs designed to mitigate the centuries-old caste system in India. They find that Affirmative Action programs are indeed effective and have limited adverse effects in the short-term for potentially underprepared treated participants.

The paper proceeds as follows: we start in Section 2 by presenting our empirical findings.

Affirmative Action

Then, we introduce our model in Section 3. In Section 4, we discuss the calibration and show the ability of our model to match non-targeted moments. We use our model to evaluate the effectiveness of Affirmative Action policies in Section 5. In this section, we also show that the channel novel to our paper provides the crucial long-term justification for short-term imposition of Affirmative Action. Last, in Section 6, we conclude and discuss avenues for future research.

2 Empirical Evidence

In this section we establish three key regularities about beliefs and their effects on labor market outcomes. First, we show that Black workers believe that discrimination will negatively affect their labor market outcomes and that these beliefs change across generations. Second, we demonstrate that hiring managers believe Black Workers experience worse labor market outcomes because they are lower ability and exert less effort on the job. We also show that, like Black workers, these beliefs evolve over time. Last, we show that Blacks who believe discrimination will negatively affect their job prospects aspire for less-prestigious occupations, that these occupations have a higher proportion of Black representation than occupations of similar prestige, and that at age 35 these workers are indeed employed in less-prestigious occupations.

2.1 Beliefs about Discrimination

We start by exploring how beliefs about discrimination differ between Black workers and supervisors who are likely to have the authority to hire workers. To do this, we use the General Social Survey (GSS), which is a cross-sectional survey conducted annually beginning in 1972 until 1994 when it became bi-annual. What distinguishes the GSS from the other surveys that span the same timeframe is that the GSS asks a variety of questions about what respondents believe cause lower incomes and worse labor market outcomes for Black Workers

than for White workers. We first focus on questions about four specific explanations asked in every survey starting from 1985. Respondents indicate whether they believe explicit discrimination, lower innate ability, lack of access to quality education, and subpar work effort contribute to observably worse outcomes for Blacks.

Beliefs about the sources of discrimination vary substantially by race, economic strata, occupation, and a number of other dimensions that provide insight into how key segments of the population view discrimination. Notably, Blacks place substantial weight on explicit discrimination as the source of differences initially, but this trends down over time. We plot these trends in Figure 2.1.

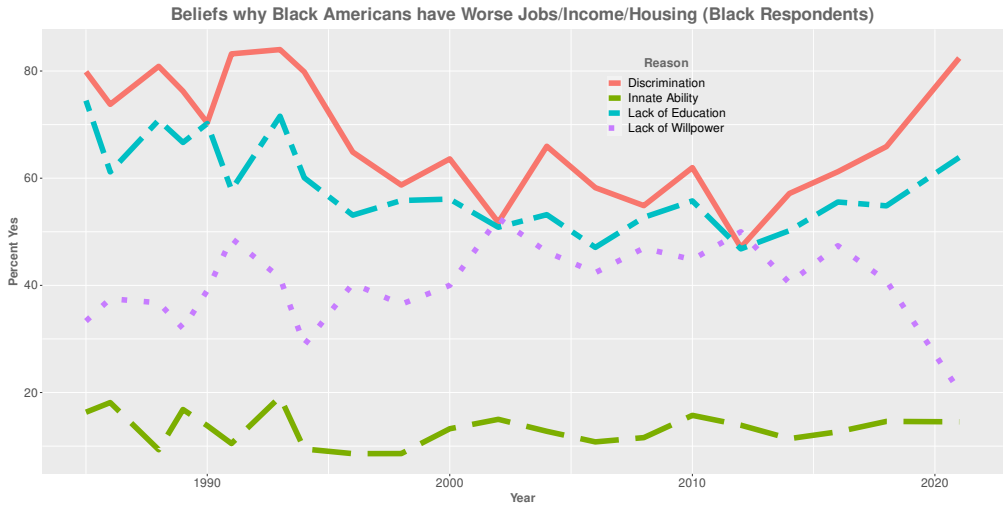


Figure 2.1: Black Beliefs about Discrimination.

By contrast, supervisors place less weight on discrimination and more weight on lack of effort as explanations for the observably worse outcomes for Blacks in the United States. We plot these trends in Figure 2.2.

While about 20% of both Black respondents and respondents who supervise workers indicate that differences in innate ability explains the differences in outcomes between Whites and Blacks, the difference lies in beliefs about willpower. Around 60% of supervisors in the 1970s and 1980s blame a lack of willpower on the part of Black workers as a reason for worse outcomes, while less than 40% of Black respondents share this belief. These findings

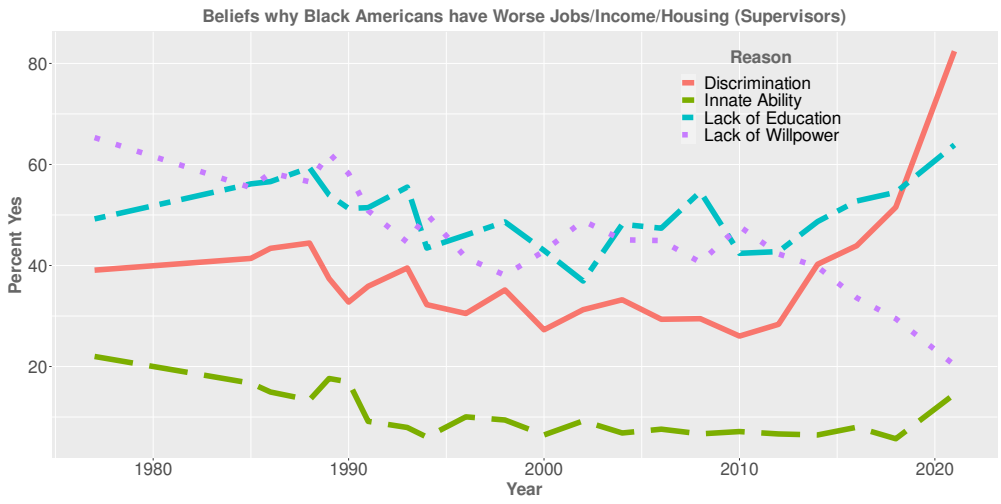
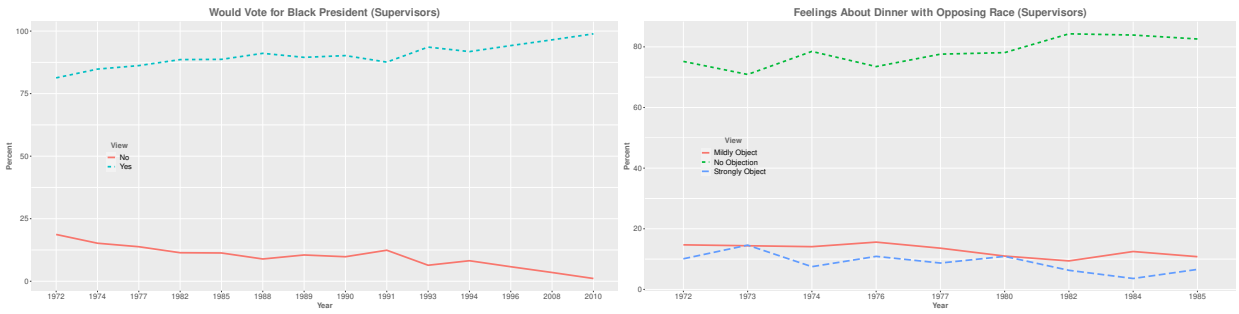


Figure 2.2: Supervisor Beliefs about Discrimination.

suggest that supervisors hold discriminatory beliefs, but do not necessarily provide insight into whether explicit taste-based racial discrimination drives these responses.

Further exploration of supervisor responses to other questions suggest a sizable share have a preference for White individuals over Black individuals. One informative question in the GSS asks respondents whether they would vote for a Black Presidential Candidate who shares their policy views, thus making race the only source of variation. The GSS also asks those surveyed whether they would feel comfortable having dinner with an individual of the opposite race. We plot the responses of supervisors over time in Figure 2.3. In the left panel we plot the share who would vote for a Black presidential candidate with the red line corresponding to a "no" response and the dashed blue line indicating "yes." In the right panel we plot the level of comfort having a person of the opposing race over for dinner, with the dotted green line indicating "no objection," the solid red line being "mild objection," and the dashed blue line indicating "strong objection."

Around 20% of supervisors respond to these questions in ways that suggest an explicit preference for White individuals for reasons outside beliefs about their productivity. Next we explore how beliefs about discrimination affect the career decisions of Black workers.



(a) Would vote for Black presidential candidate.

(b) Level of comfort having opposite race for dinner.

Figure 2.3: Further supervisor responses about issues of race.

2.2 Discrimination Beliefs and Occupational Choice

Our previous section shows that a sizable share of Blacks believe discrimination affects or affected their labor market experience. In this section we provide evidence that beliefs about labor market discrimination changes human capital investment and alters occupational choices.

We use the 1979 National Longitudinal Survey of Youth (NLSY79), which asks respondents between ages 14 and 22 whether they believe discrimination will affect their future labor market. It also asks which occupation they intend to hold at age 35. To assess whether discrimination affects occupation choice, we consider two measures of "occupational prestige." First, we merge in an occupational prestige index that ranks each youths aspired occupation by its "social standing," constructed by Duncan (1961). Second, consider the average income of an individuals aspired occupation, averaged between 1972 and 1982.

Using this data, we run the following empirical specification:

$$\begin{aligned} PrestigeIndex_i = \beta_0 + \beta_1 \times \mathbb{1}_{Black} + \beta_2 \times \mathbb{1}_{Belief} + \beta_3 \times \mathbb{1}_{Black} \times \mathbb{1}_{Belief} \\ + \beta_4 AFQT + \delta \times X_i + \epsilon_i \end{aligned} \tag{2.1}$$

where we include indicators for the education, region of residence (the most granular measure of geography in the public NLSY79) and age at which an individual, i, responds to the

survey, as well as the highest grade completed by the respondent's father and an indicator for their father's primary occupation in X_i. Last, we include the respondent's Armed Forces Qualifying Test (AFQT) percentile as a proxy for innate ability. We repeat the same specification using the real (1999$s) average income of the aspired occupation calculated at the 3-digit level using the Current Population Survey (CPS), averaged between 1972 and 1982,[2] as the dependent variable. For both estimations, we restrict our sample to Black and White male respondents. We present out findings in Table 2.1.

	Prestige Score of Career Aspiration in 1979	Aspired Occupational Income, 1970s
Black	7.0448**	4768.1849
	1.4598	2182.1669
Believes Discrimination will Affect Career=1	7.34**	13358.49***
	2.29	1681.90
Black × Believes Discrimination will Affect Career=1	-8.93***	-16090.41**
	1.35	4652.43
Test: $H_0: \beta_1 + \beta_3 = 0$	-1.881***	-11322.22**
SE	0.260	2483.498
Observations	1296	1164

Clustered standard errors in parentheses
* p< 0.1, ** p< 0.05, *** p< 0.01

Table 2.1: Columns 1 estimates the effects of discrimination beliefs on the prestige of an individual's aspirational occupation. Column 2 repeats this exercise with average income of the aspired occupation between 1972 and 1982.

To determine if discrimination beliefs negatively affect Black youths, we test whether or not $\beta_1 + \beta_3$ can be statistically distinguished from zero. Our test shows that the sum is indeed significantly negative, indicating that Black youths who hold these beliefs aspire to occupations that are both less prestigious and offer lower pay on average. Because this question is asked prior to entering the labor market and generally before an individual commits to an occupation, this is a strong indication that beliefs about discrimination do indeed affect the occupational choice of Black youths.

Next, we show that Black respondents who report a fear that discrimination will affect

[2]We use this timeframe both because individuals may form beliefs earlier than ages 14 to 22, which are likely to persist and because this is the period in which the CPS uses the same 1970 occupation codes as the NLSY79

their labor markets asipre to occupations with larger Black representation and that these aspirations are realized by age-35. We calculate the share of an occupation that is Black again at the 3-digit level using the CPS from 1972 to 1982. We merge this data with the occupational aspirations from respondents in 1979. Then we run the same specification as Equation 2.1 first with proportion of Black workers in an aspirational occupation as the dependent variable, and then with the prestige of an individuals age-35 occupation. We include a specification with (right) and without (left) father's occupation to increase our sample size. We present our results in Table 2.2

	Percent of Black Workers in Aspired Career	Prestige of Age-35 Occupation
Black	-0.0063	-2.4247
	0.0031	1.3379
Believes Discrimination will Affect Career=1	-0.01***	2.09
	0.00	3.45
Black × Believes Discrimination will Affect Career=1	0.03**	-4.99
	0.01	2.26
Test: $H_0: \beta_1 + \beta_3 = 0$	.024*	-7.419***
SE	0.010	1.178
Observations	1164	1293

Clustered standard errors in parentheses
* p< 0.1, ** p< 0.05, *** p< 0.01

Table 2.2: Percent of workers who are Black in aspired occupation.

Although not as strong as our prestige effects findings, we find evidence that Black workers are more likely to aspire to careers with a higher proportion of Black workers when they fear the effects of discrimination on their career. We also find strong evidence that this is indeed realized at age-35.

Taken together, our findings suggest that discrimination plays an important role in determinining the labor market choices of Blacks. When they anticipate limited prospects due to discriminatory practices, Blacks forgo the costly investment of prestigious careers. Instead, they enter careers where previous employment statistics suggest that Blacks are welcome. We use these findings as well as our findings in Section 2.1 as inspiration for the construction of our model in the next section and as key moments to help discipline it in our calibration,

Section 4.

3 The Model

3.1 Environment

In our environment, time is discrete and continues forever although workers exit the model deterministically at the end of the period. There are two types of firms: a fixed number of "high-prestige" firms that offer high paying jobs whose pay depends on worker productivity, and a continuum of "humble" firms that offer a fixed pay that is independent of worker characteristics. There are a fixed number of workers, who may be either Black or white ($r \in \{b, w\}$), and differ by their innate ability, $q \sim Q(\cdot)$, which we assume does not differ by race. High-prestige firms open vacancies to which workers may apply. When applying for these vacancies, workers understand that employment is probabilistic, but can be influenced after applying by the firm's preferences and a signal they receive about the worker's human capital investment.

Workers maximize their likelihood of employment and their return to becoming employed, while accounting for the cost associated with acquiring the skills necessary. After entering the model, and learning their initial conditions (r, q), workers choose to invest in human capital,[3] z, at a utility cost $C(z|q)$. In expectation, the return on their investment is given by $\Pr(e|z, r; \hat{\eta})\, z$, composed of their probability of employment given human capital and race and beliefs about discrimination, $\hat{\eta}$, times their human capital.

Following their investment, workers meet a high-prestige firm at random and choose whether or not to apply for the job. They compete for this job with $H \sim \Omega(\cdot)$ other potential applicants (whose identities are unknown). On average, this stochastic process yields H_A competitors for the job, though not all potential applicants may choose to apply.

[3] Because decisions in our model are static, there is no distinction between occupation-specific and general human capital.

Firms then score applicants and hire the applicant with the highest score. Race may affect this hiring decision because the firm observes a noisy signal about the workers productivity, similar to Phelps (1972), after receiving a workers application and also because some firms hold a taste-based preference for White workers, similar to Becker (1957). Each firm filters this signal using their beliefs about worker productivity, $f(z|r)$, which is firm-specific, varies by race, and is based on previous signals and hiring decisions. Upon hiring a worker, the firm learns their true productivity and Bayesian updates their beliefs about worker productivity using information from the productivity of the hired worker and the signals of workers who were not hired.

A measure μ of firms have an explicit taste-based preference for white workers. This affects the probability of employment for both Black and White workers. When $\mu > 0$, fewer Black workers will be hired for an occupation even if human capital is identical between Black and White applicants and firms hold identical beliefs about productivity. Both Black and White workers understand that some firms are discriminatory, but do not know the true measure, μ. Instead, they construct beliefs about discrimination by observing labor market outcomes. Because workers do not observe individual applications or the hiring history of individual firms, they are unable to distinguish between statistical and taste-based discrimination. Instead, they estimate a reduced-form measure of discrimination, $\hat{\eta}$ based on the observed employment outcomes and the resolution to the worker's investment decisions by race. These beliefs update over time as new generations observe the labor market outcomes of the previous generation.

We divide the discussion of the model into two sections: we first describe the static optimization problem faced by the current generation of workers and firms. In the subsequent section, we describe the updating problems faced by firms and workers in the following generation. We conclude by defining the equilibrium.

3.2 Static Problem

3.2.1 Worker's Problem

Workers enter the period characterized by their race, $r \in \{w, b\}$, their innate ability, $q \sim Q(\cdot)$, and beliefs about the measure of discrimination in the labor market, $\hat{\eta}$, based on the labor market outcomes of previous generations. Workers make three decisions: first they make a human capital investment decision; then, they choose whether or not to apply to a "high prestige" job, whose employment probability depends on a noisy signal of their human capital. Should they not apply or fail to be hired, they may gain employment probabilistically at a "humble" job, which pays a fixed rate z_L. Obtaining human capital incurs a cost, $C(z|q)$, which is increasing and convex in z and decreasing in q.

Given this initial state, they calculate the probability of employment given a level of investment and then choose to invest.

$$V_I(z, r; \hat{\eta}) = \max_z \{E[V_A(z, r, \hat{P}, \hat{\eta}; \nu)] - C(z|q)\} \tag{3.1}$$

In this value function, the expectation is over a application cost shock, $\nu \sim$ Gumbel, which may be positive, but is on average negative.This ensures that the model does not yield degenerate beliefs because in every cohort there is a positive probability of a worker applying and receiving a job offer. After investing, workers observe the cost of applying for the job, ν, and choose whether or not to apply for a high prestige job. Given the utility shock, workers solve the following:

$$V_A(z, r, \hat{\eta}; \nu) = \max \left\{ \hat{P}(e|z, r; \hat{\eta})\, z + (1 - \hat{P}(e|z, r; \hat{\eta}))V_L + \nu, V_L \right\} \tag{3.2}$$

where V_L is employment in the humble occupation and yields identical value for each worker. A worker with human capital z has a probability $\Pr(e|z, r; \hat{\eta})$ of obtaining a high prestige

job, in which case the earn $z > z_L$ at the end of the period. With complementary probability, they are unable to find such a job. Should they choose not to apply or are unable to find a job, they enter the market for humble jobs, which yields the following value:

$$V_L = \Pr(e_L)\, z_L + (1 - \Pr(e_L))b \tag{3.3}$$

where $\Pr(e_L)$ is a fixed probability of employment, z_L is a fixed payment that does not depend on human capital, and b is unemployment utility should they fail to be hired in this "humble" pay market.

3.2.2 Firm's Problem

Firms operate two types of technologies: a "high-prestige" firm produces linearly based on their worker's productivity, while the "humble" firm produces an identical amount, z_L in every match. Because output at high-prestige firms depends on worker characteristics, firms operating this technology attempt to hire the most productive worker from their pool of applications. A firm receives a noisy signal $y = ln(z) + \epsilon$ for each applicant, where ϵ is a normally-distributed white noise variable with variance σ_ϵ^2, following Phelps (1972). The firm then constructs a score for each applicant, based on (i) the expected value of z given y and r; (ii) a discrimination indicator $\mathbb{1}_D^B \in \{0, 1\}$; and (iii) an Affirmative Action policy indicator $\mathbb{1}_{AA}^B \in \{0, 1\}$ that affects how race enters the score. Both the discrimination indicator and the Affirmative Action indicator are always 0 for White applicants and may take a value of 1 for Black applicants depending on the firm's preferences and the policy regime. The score takes the following form:

$$\begin{aligned} s(y, r) &= \mathbb{1}_D^B \left(E[\widetilde{z}|y, r] - \gamma + \mathbb{1}_{AA}^B \zeta \right) + \left(1 - \mathbb{1}_D^B\right) \left(E[\widetilde{z}|y, r] + \mathbb{1}_{AA}^B \zeta \right) \\ &= E[\widetilde{z}|y, r] - \mathbb{1}_D^B \gamma + \mathbb{1}_{AA}^B \zeta. \end{aligned}$$

where $E[\widetilde{z}|y,r]$ is the expected z given the signal, y. The underlying conditional distribution, $f(z|r)$ may result in biased predictions by race. To interpret this function, suppose $\mathbb{1}_{AA}^{B}=0$; then a firm with $\mathbb{1}_{D}^{B}=1$ will reduce the score of a Black household by γ relative to a White household. The presence of an Affirmative Action program ($\mathbb{1}_{AA}^{B}=1$) provides incentives to hire Black applicants by downweighting the expected productivity in the score; ζ is a "bonus" productivity applied under Affirmative Action, though it does not affect production or pay. Note that we denote the productivity that the firm believes the worker to have by $\widetilde{z}$; the firm computes the expected productivity of an applicant based on the signal and race. We then suppose the firm simply hires the applicant with the highest value of s, their score.

Firms use their beliefs over productivity by race to form expectations over productivity. They do this by using the observables (y,r), as follows:

$$E\left[\widetilde{z}|y,r\right]=\int\widetilde{z}f\left(\widetilde{z}|y,r\right)d\widetilde{z}$$

To calculate this expression, they use Bayes' rule along with their priors over productivity by race, $f(z|r)$, to find

$$E\left[\widetilde{z}|y,r\right]=\int\widetilde{z}\frac{f\left(y|\widetilde{z},r\right)f\left(\widetilde{z}|r\right)}{\int f\left(y|\widehat{z},r\right)f\left(\widehat{z}|r\right)d\widehat{z}}d\widetilde{z}. \tag{3.4}$$

This expression yields the expected value of a given signal by race. The degree of optimism or pessimism by each firm affects $f(z|r)$, which changes the degree to which their beliefs are biased when presented with a signal. We discuss the formation of this belief in Section 3.3.2.

After hiring, the firm observes the hired worker's "true" productivity. This affects the beliefs about the human capital distribution by race of applicants of next period firms offering the same occupation. Simultaneously, a new generation of workers enters the labor market and updates their beliefs about the measure of discriminatory firms based on observed hiring by race. The resolution to the labor market leads workers and firms to update their beliefs,

upon which we expound in the next section.

3.3 Updating Beliefs

3.3.1 Worker Beliefs

Workers are assumed to observe the history of aggregate employment outcomes by race, $\{(e_0^B, e_0^w), ..., (e_{t-1}^B, e_{t-1}^w)\}$, but are assumed to not observe the realized application and investment decisions of individual workers, or the signal, score, and hiring decisions of individual firms. As a result, they cannot distinguish between taste-based and statistical discrimination and assume that any underperformance in the labor market is due to taste-based discrimination. Workers know the decision rules of previous cohorts (but not individual outcomes) and predict employment outcomes by supposing that firm beliefs are unbiased. This yields a set of predicted employment probabilities $\{\tilde{P}_0(e|r;\hat{\eta}_0), ..., \tilde{P}_{t-1}(e|r;\hat{\eta}_{t-1})\}$. They adjust their beliefs about discrimination in the labor market, $\hat{\eta}$ to minimize the difference between what they expect for employment outcomes and what they observe, using the history of employment outcomes.

$$\min_{\hat{\eta}} \sum_{j=0}^{t-1} \left(\frac{\tilde{P}_j(e|B;\hat{\eta})}{\tilde{P}_j(e|B;\hat{\eta}) + \tilde{P}_j(e|W;\hat{\eta})} - \frac{e_j^B}{e_j^B + e_j^W}\right)^2 \tag{3.5}$$

where $\tilde{P}_j(e|B;\hat{\eta})$ is given by

$$\hat{\eta}\tilde{P}_j(e|B,D;\hat{\eta}_j) + (1-\hat{\eta})\tilde{P}(e|B,N;\hat{\eta}_j) \tag{3.6}$$

With an updated guess $\hat{\eta}$, workers are able to predict their employment probability for a given level of human capital investment, z. Without loss of generality, we consider the calculation conducted by worker 1. To compute the probability of employment given (z, r),

workers compute the probability that worker 1 receives a higher score than workers $\{2, ..., H\}$, formally

$$\Pr\left(s_1 \geq \{s_2, ..., s_H\}\right),$$

which is the first order statistic. This probability depends on the productivities chosen by the other applicants; we suppose that these choices are made using a decision rule $Z^0(q, r)$. Because workers are unable to distinguish between discriminatory firms and non-discriminatory firms, they construct a score $\hat{s}$ using their guess $\hat{\eta}$ of the likelihood that any firm is discriminatory. This yields a score

$$\begin{aligned}\widehat{s(z,r)} &= \hat{\eta}\left(E\left[z|y,r\right] - \mathbb{1}_D^B\gamma + \mathbb{1}_{AA}^B\zeta\right) + (1-\hat{\eta})\left(E\left[z|y,r\right] + \mathbb{1}_{AA}^B\zeta\right) \\ &= E\left[z|y,r\right] - \hat{\eta}\mathbb{1}_D^B\gamma + \mathbb{1}_{AA}^B\zeta.\end{aligned}$$

for a given worker with z level of investment. Because workers are unable to distinguish between taste-based and statistical discrimination, $E\left[z|y,r\right] = ln(z) + E[\epsilon] = z$, their chosen level of human capital investment.

Workers use this calculated score to determine their probability of employment for any level of investment, z. Given the decision rules of other agents, Z^0, and their beliefs about discrimination, $\hat{\eta}$, they compute the probability of employment as

$$\begin{aligned}\Pr(e|z,r;Z^0,\hat{\eta}) = \prod_{t=2}^{T}[\int_{\epsilon_t}\int_{q_t}\sum_{r_t=0}^{1}(\int_{s(Z^0(q_t,r_t)+\epsilon_t,r_t)}^{s_{\max}}(\frac{1}{\sqrt{2\pi}\sigma_\epsilon}\exp(-(y(s,r)-ln(z))^2 2\sigma_\epsilon^2))|\frac{\partial y}{\partial s(y,r)}|ds) \\ \times \frac{1}{\sqrt{2\pi}\sigma_\epsilon}\exp(-\frac{\epsilon_t^2}{2\sigma_\epsilon^2})Q(q_t)p(r_t)dq_t d\epsilon_t]. \qquad (3.7)\end{aligned}$$

This expression states that the worker computes the probability that his score (which is a function of the unknown signal y) will be higher than every other applicant for each value

of (q, r, ϵ) they might draw, which is the domain of s above the value $s\left(Z^0\left(q_t, r_t\right) + \epsilon_t, r_t\right)$, then averages over the values of (q, r, ϵ). As we discuss in Section 3.4, workers update their decision rules taking this probability as given until Z^0 converges.

3.3.2 Firm Beliefs

Firms update their beliefs based on their interviews and hiring during the previous period. They are unable to observe the hiring decisions of other firms, and thus are endowed with the following information: a set $\{\vec{y}, \vec{r}\} = \{y, r\}_{i=2}^{\hat{H}}$ of signals with the applicants corresponding race, and (y, z, r), the signal, true productivity, and race of the worker that they hired. We let $\hat{H} \leq H$ denote the number of workers who apply, because some potential applicants may choose not to pay the ν utility cost. This yields a sample given by $X = \{\{\vec{y}, \vec{r}\}, (y, z, r)\}$. Firms revise their own beliefs by constructing a likelihood of observing this sample given their priors.

Firms construct this likelihood by first acknowledging that signals are white noise. Given any productivity, race pair (z, r), they calculate the probability of receiving signal y according to the following

(z, r):

$$p\left(y|z, r\right) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon} \exp\left(-\frac{\left(y - ln(z)\right)^2}{2\sigma_\epsilon^2}\right),$$

which is the just the probability that $\epsilon = y - ln(z)$. The probability of a given signal for an applicant of race r is then

$$p\left(y|r\right) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon} \exp\left(-\frac{\left(y - ln(z)\right)^2}{2\sigma_\epsilon^2}\right) f\left(z|r\right).$$

Given priors $f(z|r)$, this is the likelihood that a firm would observe a particular signal y by race. After hiring a candidate, the firm learns that candidate's true productivity, z, and retains information on the signal received from the other $\hat{H} - 1$ applicants. Now the

firm is able to update their beliefs over the productivity distribution using this sample and their prior beliefs over productivity. We suppose this prior belief takes the form of a Beta distribution with parameters $\{\alpha, \beta\}$. This set of information yields the following likelihood of receiving a sample $X = \{\{\vec{y}, \vec{r}\}, (y, z, r)\}$ for a given (α, β) pair:

$$p(X|\alpha,\beta) = \Pi_{i=1}^{H} \left(\frac{1}{2\sqrt{\pi\sigma}} e^{-\frac{1}{2}} (\frac{y_i - ln(z)}{\sigma})^2 f(z|r,\alpha,\beta) \right)^{1-\mathbb{1}_{s_i=\hat{s}}} f(z|r,\alpha,\beta)^{\mathbb{1}_{s_i=\hat{s}}} \tag{3.8}$$

Firms integrate this expression using their prior uncertainty $g(\alpha, \beta)$ over the parameters (α, β), which yields the following probability of observing sample X:

$$p(X) = \int \int p(\{y, r\}, (y, z, r), \alpha, \beta)\, g(\alpha, \beta)\, d\alpha d\beta \tag{3.9}$$

Finally, firms apply Bayes' rule to update their posterior over the Beta parameters

$$g(\alpha, \beta|X) = \frac{p(X|alpha, \beta)g(\alpha, \beta)}{p(X)} \tag{3.10}$$

and use this posterior to field beliefs about productivity by race,

$$f(z|r) = \int \int f(z|r, \alpha, \beta)g(\alpha, \beta|X)d\alpha d\beta \tag{3.11}$$

to which they apply Bayes rule to elicit $f(z|y, r)$ and then the conditional expectation of productivity by race according to Equation 3.4. We assume that each firm begins with uniform priors over each parameter in the Beta distribution but allow this to change while solving for the fixed point that supports initial beliefs $\hat{\eta}_0$.

3.4 Equilibrium

We have a Nash equilibrium of the game if the decision rule used by the worker coincides with $Z^0(q, r)$; that is, the worker correctly anticipates the decision rules that will be used by his competitors. This equilibrium is conditional on their beliefs about discrimination, $\hat{\eta}$.

We compute the equilibrium decision rule iteratively – we guess Z^0, solve the worker problem for Z^1, and repeat until it converges. To compute the integrals, we exploit Gauss-Hermite quadrature for the normal variables $(q,\ y|z, r)$ and Gauss-Jacobi quadrature for $f(z|r)$. We use a spline approximation to $s(y)$ computed by solving the nonlinear equation at different values of y and interpolating between them to compute $y(s, r)$ when necessary (including to compute the Jacobian term).

4 Calibration

In this section, we describe our calibration approach. We first discuss preset parameters, functions, and distributions, and then describe our calibration targets and results. We conclude this section by showing that the model is able to replicate key non-targeted moments.

4.1 Empirical Preliminaries

4.1.1 Functional Form and Distributional Assumptions

We first discuss our distributional and functional form assumptions. As we discussed in the previous section, we assume that workers have linear utility, $u = z$ and face a cost function $C(z, q) = \frac{z^2}{2q}$. Workers are endowed at birth with innate ability q, which we assume is log-normally distributed, $q \sim LN(\mu_Q, \sigma_Q)$, where the parameters are assumed to be identical for both White and Black workers. Potential applicants arrive at high-prestige firms according to $H \sim exp(H_A)$, an exponential distribution with mean H_A. Workers must pay a utility cost when applying for high-prestige jobs, $\nu \sim G(\sigma_\nu)$, which we assume is Extreme Value

(Gumbel) distributed. Upon applying to a high-prestige job, the firm observes a noisy signal of the worker's productivity, $y = log(z) + \epsilon$, where we assume that ϵ is normally distributed, $\epsilon \sim N(0, \sigma_\epsilon)$. We assume that firm beliefs over the equilibrium human capital distribution (post-investment) by race is Beta-distributed, $f(z|r) \sim Beta(\alpha, \beta)$. We allow beliefs about the parameters of these Beta distributions as well as the worker belief about the degree of discrimination to be non-parametrically distributed.

4.2 Preset Parameters

We preset a selection of parameters, some of which are calibrated externally and some of which are set to reasonable values and tested for sensitivity. We set the "true" measure of discriminatory firms to 22.3% ($\mu = 0.223$), the fraction of White supervisors in the GSS who reported they believed Blacks workers had lower ability than their White counterparts. This is higher than alternative measures that we considered more incontrovertible evidence of explicit racism of the respondent, like the 4.8% who responded that they would be uncomfortable having a Black person over for dinner. However, decreasing μ only serves to increase the effectiveness of Affirmative Action, so we choose this conservative value. We also use the GSS to set $\hat{\eta}$ equal to the share of Black respondents who indicated that they believed that discrimination would affect their labor market outcomes. Because this question is first surveyed in 1985 and we calibrate our model to the mid-1970s, we linearly predict the value for 1975. This yields a value of 0.738 for 1985 and 0.784 for 1975, which we use for $\hat{\eta}_0$.

We assume that there are 1000 firms ($N = 1000$) in the economy, and assume that there a maximum of 20 possible applicants for each firm.[4] We set the proportion of Black and White applicants to reflect their relative shares of the prime age male, non-institutionalized populations in 1979, which are 93% White and 7% Black. We round these to 90% and 10%

[4]We choose 1000 due to memory constraints, which may affect the rate at which workers learn about discrimination. In our simulations, the information gain of additional firms slows precipitously after 1000 firms. We require a maximum number of possible applicants to construct grids, hence the choice of $\bar{H} = 20$

for White and Black, respectively, so the pool of potential applicants are on average integers for both races and set the average number of potential applicants to $H_A = 10$. We normalize μ_Q and z_L to 1, the mean of the innate ability distribution and the pay at a humble job, respectively. We also set $b = 0.4$, approximately the replacement rate for unemployment insurance in the United States (Shimer, 2005).

4.3 Calibration Approach and Results

This leaves us with 5 parameters to estimate: σ_Q, the standard deviation of the common innate ability distribution, σ_ϵ, the standard deviation of the noise of each application, σ_ν, the standard deviation of the application cost shock, γ, the penalty applied to Black scores by taste-based discriminatory firms, and $\Pr(e_L)$, the probability of employment in the humble job market. While these parameters are jointly estimated, we sketch out our argument for ex-ante identification. We discipline σ_Q by targeting the 95-to-5 earnings ratio. Because earnings for the high prestige market equal investment and investment is likely to be increasing in q, a larger 95-to-5 ratio implies a higher value of σ_Q, all else equal. We target the Black-to-White earnings ratio to discipline σ_ϵ. All else equal, a smaller value of σ_ϵ reduces the racial income gap by decreasing the degree of statistical discrimination in the hiring process. To discipline σ_ν, we target the employment rates at the 25th and 75th percentiles of the AFQT distribution, averaged over our sample in the NLSY79. As applying becomes more costly, employment rates below the top decile of the innate ability distribution flatten out, because fewer and fewer workers choose to apply. We target the relative likelihood of employment at high-prestige firms to identify γ. γ changes the likelihood that a discriminatory firm will hire a Black worker and thus changes the employment rates in high-prestige occupations. For $\Pr(e_L)$, we target the Black and White unemployment rates because this parameter directly controls the likelihood that a worker enters unemployment. We use prime age (25-54) males who are either Black or White to construct all of our moments.

4.3.1 Results

Our estimation yields the following set of parameters:

Parameter	Value	Comment
σ_Q	1.38	SD of Innate Ability Dist.
σ_ϵ	0.816	SD of Signal Noise
σ_ν	1.38	SD of Application Taste Shock
γ	50	Taste-Based Discrimination Score Penalty
$Pr(e_L)$	0.788	Low Prestige Employment Probability
μ	0.221	Supervisor Responses about Lower Black Ability (GSS, 1977)
$\hat{\eta}_0$	0.784	Black Responses about Labor Market Discrimination (GSS, 1985)
b	0.4	Approx UI Replacement Rate (US)
z_L	1.00	Normalization
ζ	50	Assumption
μ_Q	1.00	Normalization

Table 4.1: Parameter values.

Of note is the large value of γ relative to the innate ability distribution. This indicates that taste-based discriminators are unlikely to hire any but the most qualified Black applicants. We present the fit of our estimation in Table 4.2.

Moment	Data	Model
Black-to-White Earnings Ratio	0.6273	0.6591
95/5 Earnings Ratio (Pooled)	5.8203	5.8454
Black Unemployment Rate	0.0597	0.0401
White Unemployment Rate	0.0285	0.0390
Black Employment Rate (25th AFQT Pctile)	0.9100	0.9533
Black Employment Rate (75th AFQT Pctile)	0.9600	0.9591
White Employment Rate (25th AFQT Pctile)	0.9600	0.9622
White Employment Rate (75th AFQT Pctile)	0.9900	0.9595
Ratio of Black-to-White High Prestige Employment Rates	0.5623	0.5149

Table 4.2: Targeted moments.

The results of our estimation closely parallels our selected moments. The coarseness of our labor market limits somewhat the ability of our model to exactly match certain moments, though it consistently matches the sign of Black-White differences for comparable moments if not always the magnitude. We slightly understate the Black and over state the White unemployment rates. This is because the heterogeneity in job-finding rates exists only for the "high prestige" jobs, after which both races face an equal probability of employment. While more White workers become employed in these high prestige jobs, closely paralleling the data,

we cannot fully account for differences in aggregate employment rates and employment rates for lower AFQT workers (which proxies for q).[5] Despite the relative simplicity of the model, it does a remarkable job matching both the 95th percentile to 5th percentile earnings ratio as well as the Black-to-White high prestige employment ratio, two key moments for our estimation.

4.3.2 Non-Targeted Moments

Here, we show that our model reasonably reflects key aspects of the data. We compare the results of our model to beliefs about discrimination, and earnings and employment moments over time. We show that the model roughly tracks the evolution of Black beliefs about discrimination in the data. Then, we show that it also matches earnings across cohorts. We present our comparisons in Figure 4.3.2. In the left panel, we plot the Black-to-White earnings ratio for the data (solid blue with box markers) and the model (dashed red with circular markers). In the right, we plot discrimination beliefs, using the same color and marker scheme. In both cases, we assume that the primacy of an individual cohort lasts for 5 years.

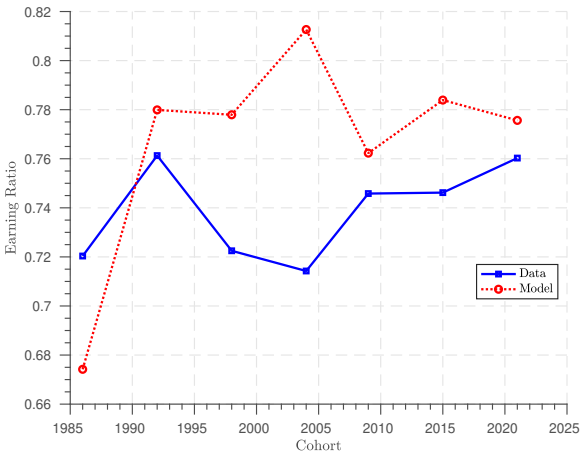


Figure 4.1: Earnings ratio (Black-to-White)

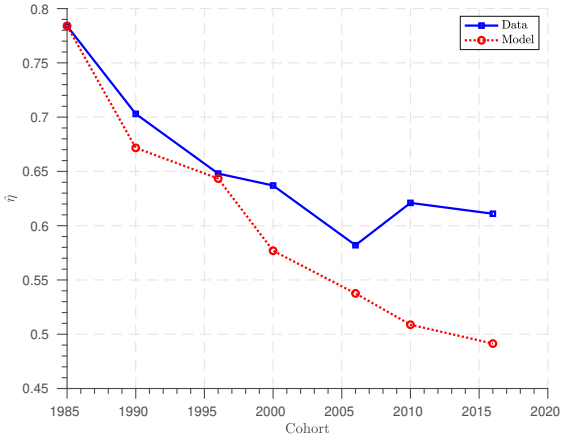


Figure 4.2: Comparison between $\hat{\eta}$ and discrimination beliefs.

[5] If labor markets were further segmented along occupations, it should be trivial to match aggregate employment measure; however, the computational burden would make the exercise untenable.

The model does a good job capturing the decline in beliefs about discrimination, though it does not yield the upward revision toward the end of the time series. We believe this is due to external factors outside the scope of the model. The model roughly captures the behavior of earnings over time same time frame, though relative earnings increase more for Black workers in the model than in the data.

5 Findings

In this section, we consider the role of beliefs on employment of Black workers. We start by exploring the interaction between beliefs, investment, and subsequently hiring for the current generation. Then, we examine how outcomes of the current generation affect future worker and firm beliefs, and propagate for many subsequent generations. Last, we impose an Affirmative Action policy and assess the resulting changes to beliefs, productivity, and relative inequality between Black and White workers. We then demonstrate that two-sided dynamic beliefs are vital for understanding the dynamics of racial inequality and for informing the debate about Affirmative Action and conclude by showing that Affirmative Action is not universally beneficial.

In each of our simulations, we hold fixed our calibrated value of $\hat{\eta}$ to find a fixed point in both $Z_r^0(q)$ and $\hat{P}(e|z,r;\hat{\eta})$. Then we conduct 20 simulations, each of which spans 20 consecutive cohorts. During each simulation, each worker in a cohort is randomly endowed with innate ability $q \sim Q$, and firms are randomly assigned to be taste-based discriminators prior to the first cohort and this designation is fixed throughout.

5.1 Model Mechanisms

Our analysis starts by exploring the mechanisms and how they interact in our baseline economy. We first explore the static optimization over human capital, applications, and hiring. Then, we show that these beliefs create feedback by altering behavior that affect

subsequent decisions.

5.1.1 Interaction between Beliefs and Human Capital Investment

We start by exploring the interaction between beliefs and the static choices made by each cohort. Every member of a cohort is tasked with a decision: how much human capital should I accrue and what costs am I willing to incur to do so? Subsequently, should I apply for a job, or forgo the (potentially) costly process and wait to likely receive "humble" job. They resolve this trade-off by assessing the likelihood that they will receive a job offer after investing, which depends on their beliefs about discrimination $\hat{\eta}$, the investment decisions of other workers, and the beliefs that firms use to filter the signals about quality that they receive from an applicant.

Because they face different labor market prospects, Black and White workers' resolutions to this trade-off can differ substantially. Human capital investment depends on a workers beliefs about their likelihood of employment for a given level of human capital. When employment prospects appear promising, a worker is more willing to undertake costly investment because the expected payout is higher, and vice-versa. Because of long-observed divergence in hiring in the labor market, White workers perceive a higher likelihood of employment for any given level of human capital, which causes them to invest more. We show these differences for the first cohort and Black and White workers in Figure 5.1.1. In the left panel, we show beliefs about the likelihood of employment given a level of human capital, z, for Black (solid blue line) and White (dotted red line) workers. In the right panel, we show the human capital investment (z) decisions of the first cohort given their innate ability, q, for Black (blue star markers) and White (Red dot markers) workers. The differences in beliefs about employment prospects in the left panel directly translates into the differences in investment decisions in the right panel between Black and White workers. Black workers with innate ability less than 50, roughly the 98th percentile of the innate ability distribution, choose not to invest in human capital. By contrast, their White counterparts begin investing in human

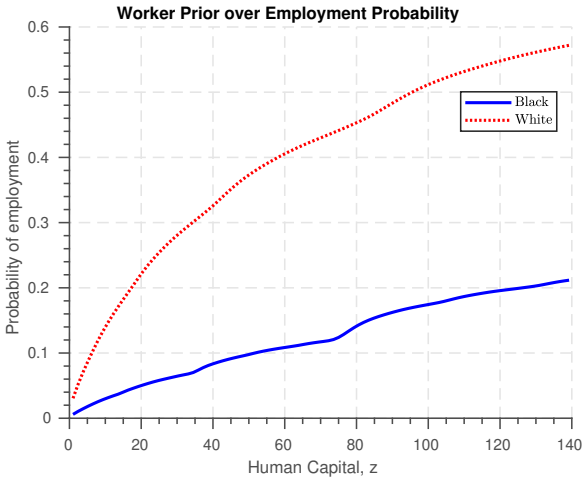


Figure 5.1: Beliefs about employment probabilities given z by race.

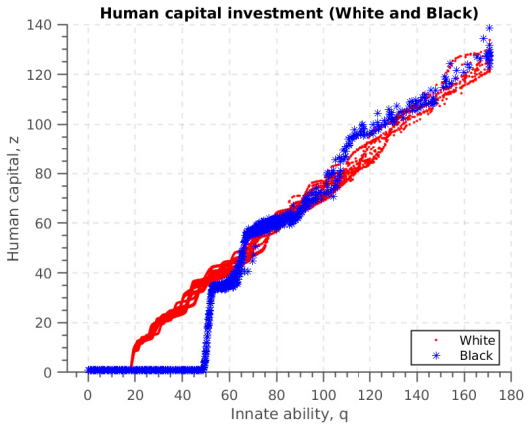


Figure 5.2: Human capital investment decisions (z) by innate ability (q) and race.

capital at lower levels of innate ability (roughly the 92nd percentile), and continue to do so up to the highest levels of innate ability.[6] As a result, there are fewer qualified Black workers for high prestige positions than their equally capable White peers.

This has ramifications for their application decisions. The bulk of Black workers forgo applying to high prestige occupations because the application process is costly and their lower investment in human capital means prospects are limited. For the few that invest or for whom applying is not costly (a positive Gumbel shock), they are met with stiff competition from White workers. In Figure 5.3, we plot the application decisions of Black and White workers across levels of human capital investment. The solid blue bars correspond to Black workers who chose not to apply for the high prestige jobs, while the yellow bars report the same for White workers. The orange forward slash bars are for Black workers who apply, while the purple backslash bars correspond to White workers who choose to apply. All bars are the natural logarithm of the count of workers at that level of human capital. Perhaps the most striking feature of this figure is the "missing" investment for Black workers: while the number of White workers applying at every level of human capital smoothly declines, there

[6]For the most wildly capable workers, we observe a reversal: Black workers invest more than White workers. This is similar to the phenomena observed by Hsieh et al. (2019), where the most capable minority and female workers invest more than their White male counterparts in order to attain employment.

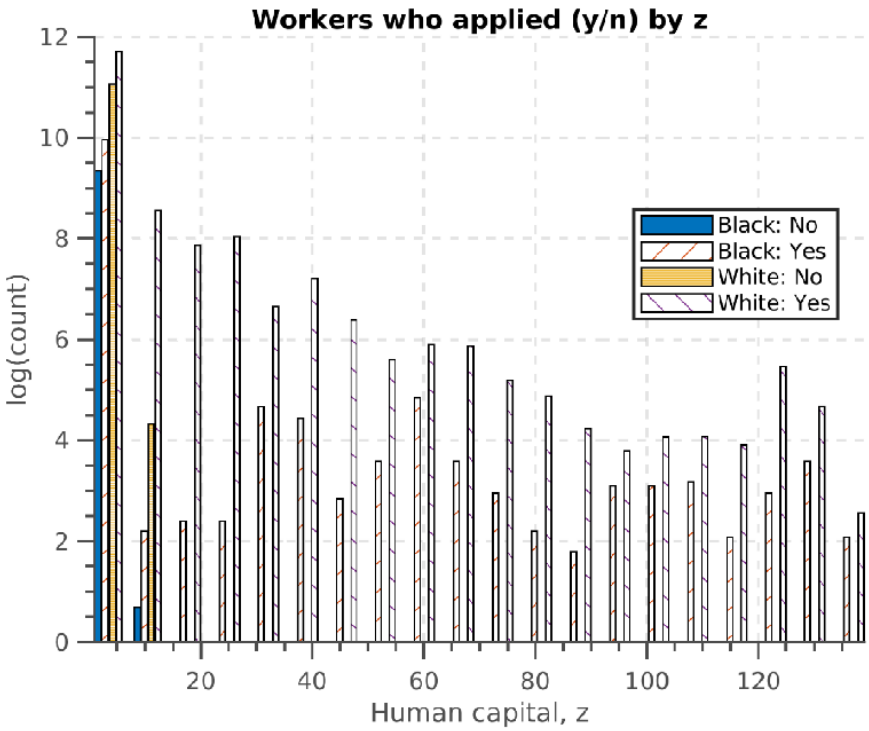


Figure 5.3: Application decisions for Black and White workers by human capital (z).

is a truncation of Black workers between the lowest levels of human capital and the point at which Figure 5.2 showed investment resumes. This has consequences for Black workers both statically and dynamically: for the present cohort, their lower application rate means that all else equal, a disproportionate (relative to their population size) number of White workers are hired. This negatively affects future cohorts because firms have fewer opportunities to interview or hire Black workers, limiting the scope within which they may adjust their priors. Additionally, a sizable share of applications by Black workers come from low-z applicants, who receive a positive ν shock, which causes firms to downweight positive signals about Black workers and affects their future hiring decisions.

The impact that these investment and application decisions have on firm beliefs is consequential. The average firm in the economy, pooling both the taste and non-taste discriminators, places more weight on low levels of z for Black workers. This means that in addition to taste-based discrimination, Black workers face endogenous statistical discrimination, in

which many firms have downward-biased beliefs about their productivity for a given signal. In Figure 5.4, we plot the average priors over human capital held by firms during the first cohort. Firm beliefs over Black productivity are the blue line, while White productivity is the red line. The mean of both distributions are given by the dashed lines with their corresponding colors, and the standard deviation of the distribution is given by the arrows emanating from the averages. This figure again shows how discrimination creates path de-

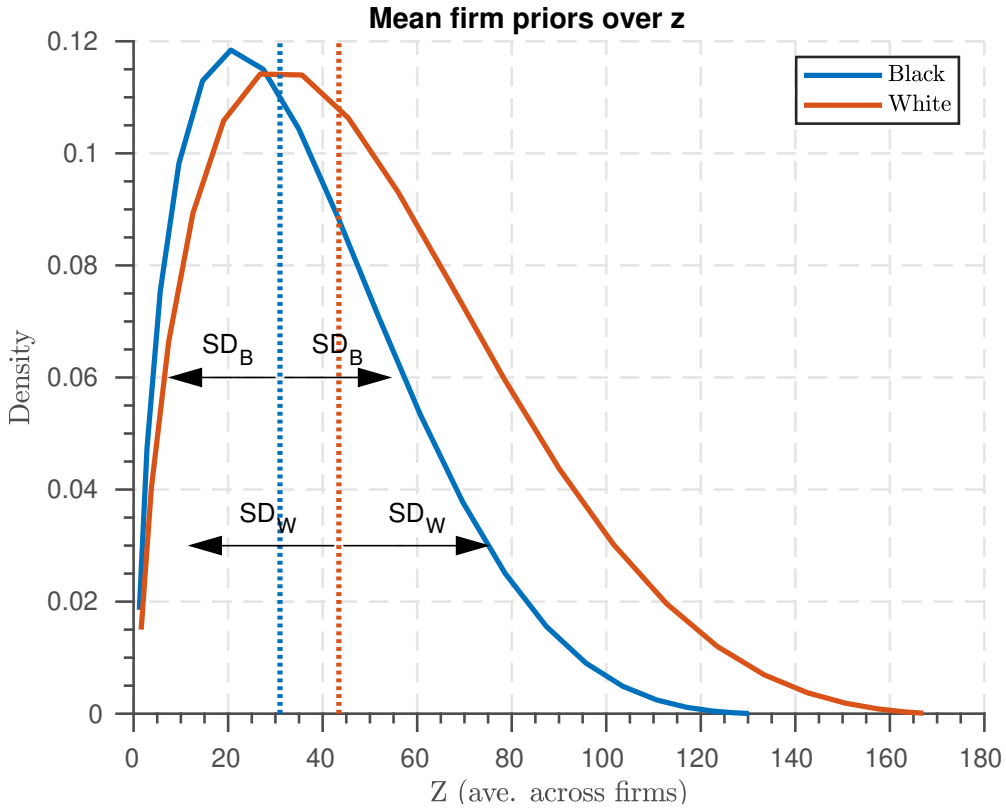


Figure 5.4: Average firm beliefs over human capital (z) by race.

pendence: generations prior to the first cohort, whose decisions and outcomes form the initial fixed point in firm beliefs invested less, which caused firms to revise their beliefs downward. Crucially, this means that for a given level of human capital, statistical discrimination is worse for Black workers.

The beliefs held by individual firms translate directly into their hiring decisions. A firm that places more weight on the probability that a particular signal is high-z is willing to hire

a worker with a worse signal, and vice versa for a firm that places less weight. In Figure 5.1.1, we plot the priors of two groups of firms: those who hired a Black worker (left panel) and those who did not hire a Black worker (right panel). In each figure, we plot the beliefs of firms who interviewed (or hired) a "high-z" Black worker in blue, which corresponds to the top percentile of the Black z distribution, and a "low-z" Black worker in dashed red, which corresponds to bottom percentile of human capital investment.

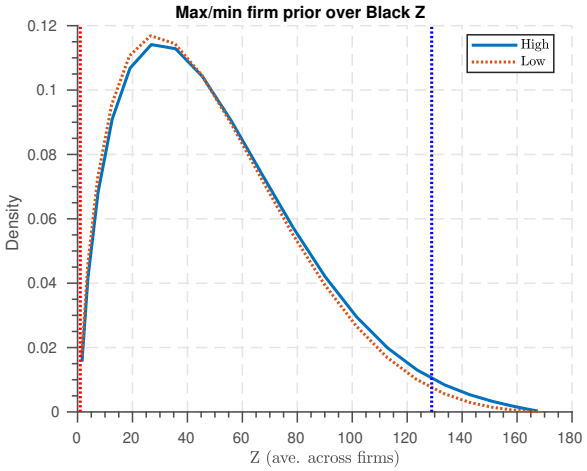


Figure 5.5: Hired Black worker

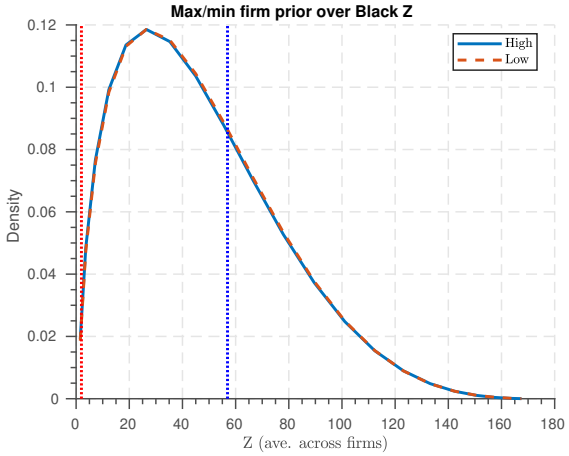


Figure 5.6: Did not hire Black worker

These figures demonstrate the importance of firm beliefs. In the left panel, the firm that hired the "high-z" Black worker has, on average, less optimistic beliefs about Black productivity than the firm that hired the "low-z" Black worker. It took an application from the top percentile of the Black human capital distribution to induce them to hire the applicant. By contrast the firm that hired the "low-z" Black worker placed less weight on lower values of z, leading then to filter the signal more optimistically.

At first glance, the right panel of Figure 5.1.1 appears puzzling. The priors of the firm that interviewed the "high-z" Black worker look similar to the "low-z" hiring firm in the left panel; the difference is that for a firm not to hire a Black worker from the top of the human capital distribution, they must have received a very bad signal (i.e., $\epsilon << 0$), or be a taste-based discriminator and thus unlikely to ever hire a Black worker. The "low-z"

non-hirer correctly filters the signal to indicate that this applicant has a low level of human capital and turns to an alternate candidate to hire.

Each of these figures demonstrates that accurately filtering an applicants signal is the crucial step in the hiring process. Biased beliefs may lead firms to only hire the most qualified applicants, like the "high-z" firm in the left panel, or errantly hire unqualified applicants, like the "low-z" firm in the left panel, which can have dynamic consequences for beliefs. In Figure 5.1.1, we plot $(z, E[z])$ pairs for individual applicants. In the left panel, we compare the (z, E[z]) pairs of White workers with red unfilled circles, and the (z, E[z]) pairs of Black workers with blue x marks, using a scatter plot of Black z and $E[z]$ based on their signal and the firms filtering. In the right panel, we conduct a similar exercise, comparing non-taste discriminators (blue "x" marks) and taste discriminators (red unfilled circles). In both figures we drop non applicants and show the lines of best fit, with colors corresponding to their respective group. Together, the figures reveal the extent to which firms hold statistical

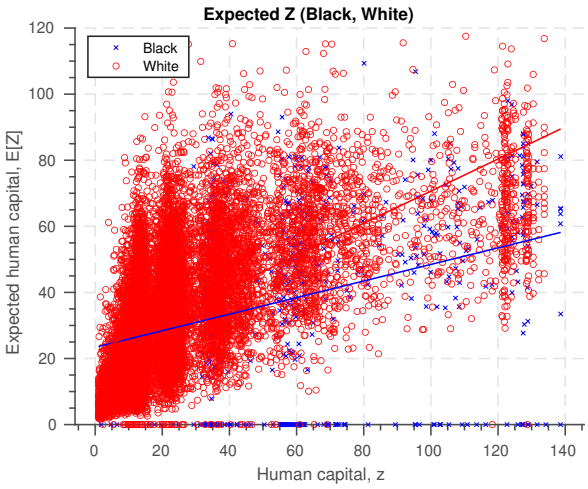


Figure 5.7: Black and White Z and E[Z]

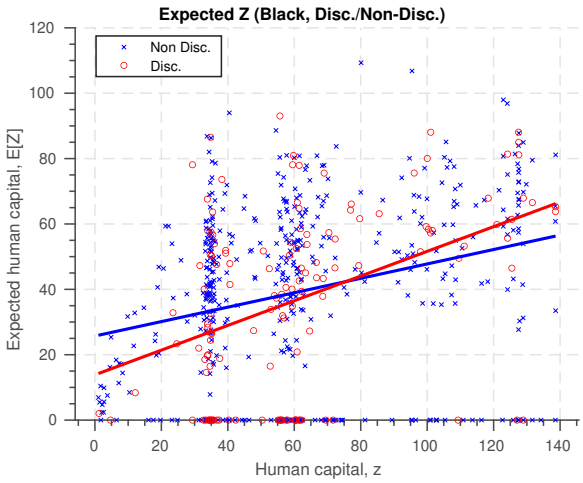


Figure 5.8: Black Z and E[Z] comparison by disc.

biases against Black workers, and a partial picture of the source of those biases. In the left panel, the line of best fit shows that the correlation between the true human capital of a Black worker, z, and the firm's filtered belief about their productivity, $E[z]$, is biased substantially downward relative to their White counterparts. The right panel provides some additional

nuance: discriminatory firms rarely hire Black workers, but when they do it is because their expected human capital is very high, leading to less biased beliefs on average–though this provides little recompense, as taste discriminatory are highly unlikely to hire Black workers.

Our analysis to this point has dealt with the decisions and consequences of agents based on the experiences of past generations. However, these static decisions play an important role for future generations that we will explore in the next section.

5.1.2 Propagation through Future Beliefs

Beliefs are the exclusive driver of dynamics in the model. Each generation takes as given the labor market outcomes of previous generations before determining their investment, application, and hiring decisions. The impact of the decisions made by previous decisions on these beliefs provides crucial insight into both the dynamics of the model and the role that Affirmative Action will play in the coming analysis.

Prior to choosing investment, workers update their beliefs about their labor market prospects based on the hiring outcomes of the previous generation. They base their updated views on the investment rules of previous generations and how these rules would translate into Black and White employment given a degree of discrimination in the labor market. If Black workers outperform these expected outcomes, the new generation revises their beliefs about discrimination, $\hat{\eta}$ downward; similarly, if the previous generation of Black workers underperform, the present generation revises $\hat{\eta}$ upward. They translate this belief into their key decision-making object: the probability that, given human capital z, they will obtain employment at a high prestige firm, $\hat{P}(e|z, r; \hat{\eta})$.

We show how this varies based on the previous generations labor market outcomes in Figure 5.1.2. In the left panel, we plot a host of counterfactual hiring outcomes for the first cohort. We then calculate the next generations updated beliefs about discrimination, $\hat{\eta}$ for each counterfactual outcome (red circles), as well as the true outcomes (blue asterisks). The relationship is clear: when the previous generation of Black workers achieve higher levels

of high prestige employment, the next generation reduces their beliefs about labor market discrimination. In the right panel, we plot the corresponding beliefs about the probability of employment ($\hat{P}(e|z,r;\hat{\eta})$) for the true $\hat{\eta}$ (dashed yellow line) and a higher value of $\hat{\eta}$ (solid blue line) and a lower value of $\hat{\eta}$ (dashed-dot green line). This figure again shows the importance of beliefs: higher values of $\hat{\eta}$ cause workers to be more pessimistic about their employment prospects for every level of investment and vice-versa. These figures both show that labor market outcomes can propagate to future generations, either positively or negatively.

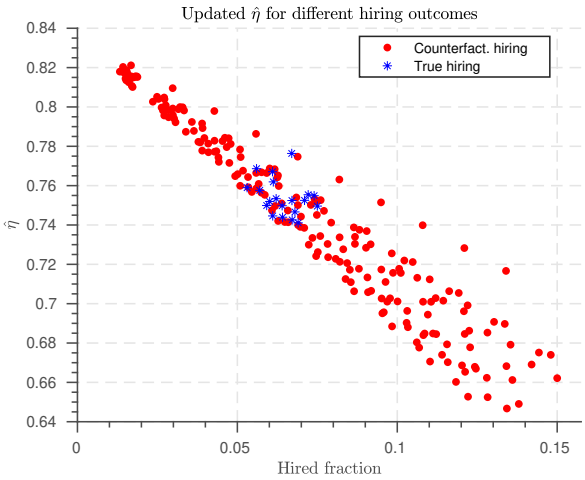


Figure 5.9: $\hat{\eta}$ outcomes for different employment outcomes.

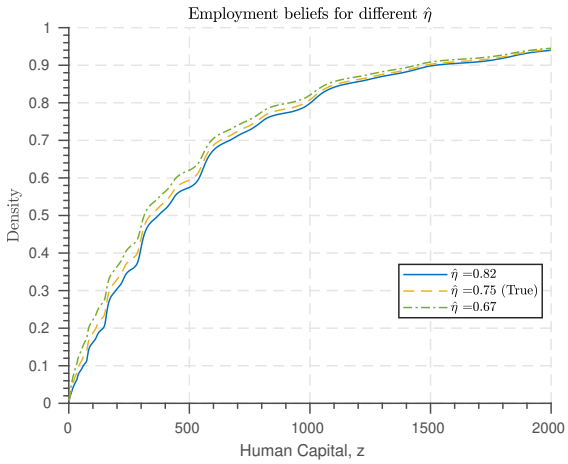


Figure 5.10: Black worker beliefs over employment probability for different values of $\hat{\eta}$.

Next, we explore the impact that hiring and interviewing workers has on firms. We return to the firms described in Figure 5.1.1 from the previous section and examine the impact on their posterior beliefs. We plot the change in resulting densities from prior to posterior in Figure 5.1.2. As before, the left panel consists of firms that hired either a "high-z" (blue) or "low-z" (red) Black worker from the first cohort, while the right panel consists of firms that interviewed, but did not hire a Black worker, again with "high-z" in blue and "low-z" in red.

These figures show a key propagation mechanism. The firm that interviewed the "low-

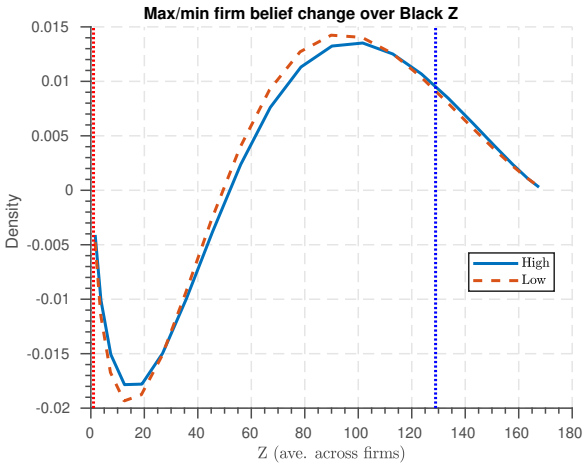


Figure 5.11: Change after hiring Black worker

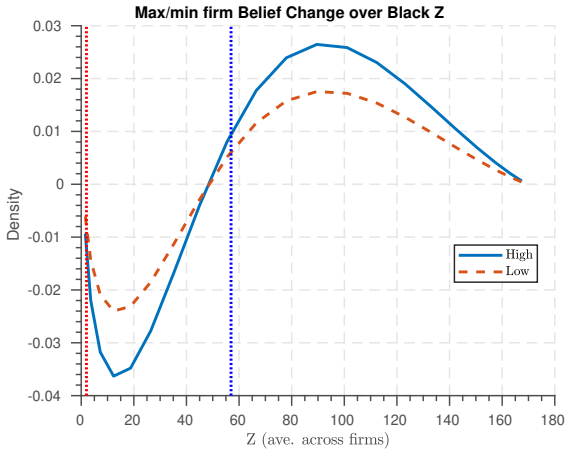


Figure 5.12: Change after not hiring Black worker

z" Black worker and believed the signal to be an underestimate downweights values of z immediately above the revealed human capital of the worker (red vertical dotted line), in order to avoid making the same mistake in the future. By contrast, the "high-z" firm increases the weight on higher values of z, even relative to the already optimistic "low-z" firm. A similar phenomena can be observed for the non-hiring firms. The firm that interviewed the "high-z" Black candidate puts more weight on higher values of human capital in the future than the firm that interviewed the "low-z" applicant.

Each subsequent cohort updates their beliefs and statically optimizes in the same manner. The sequence of hiring outcomes, consistently increasing across each cohort, yields a decline in beliefs about discrimination, $\hat{\eta}$. We plot both of these series in Figure 5.2.2. On the left, we show that these gains in employment result in decreases in beliefs about discrimination in each subsequent cohort. On the right, we show the probability that any individual Black worker gains employment at a high prestige firm.[7] As beliefs about discrimination fall, employment prospects brighten and Black workers alter their human capital investment strategies. We plot average human capital of Black (left panel) and White (right panel) workers in Figure 5.1.2.

[7]Because we allow for 20 applicants, the unconditional probability that any applicant receives a job is 1/20.

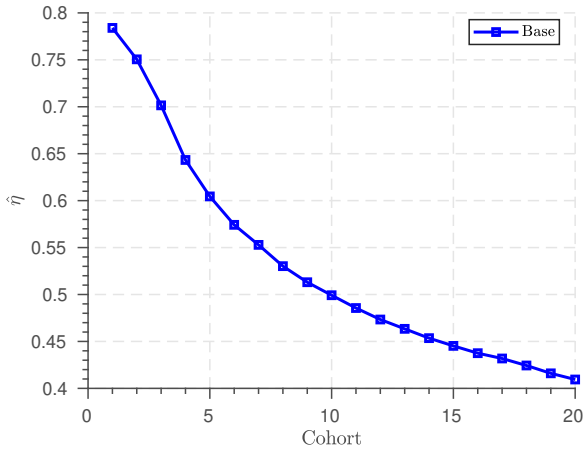


Figure 5.13: $\hat{\eta}$ across cohorts

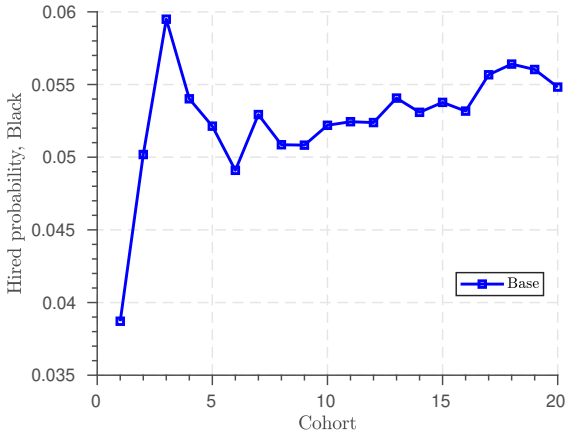


Figure 5.14: Black emp. outcomes

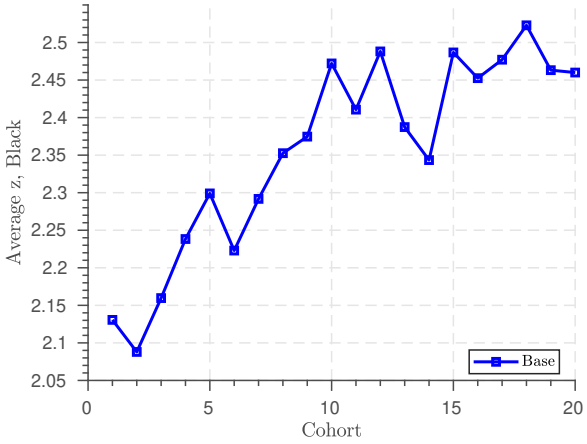


Figure 5.15: Human capital z (Black)

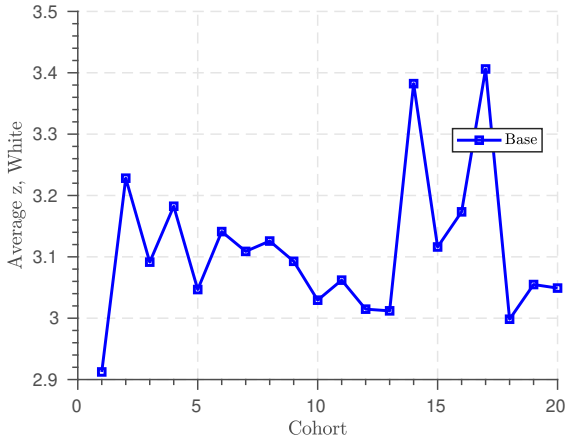


Figure 5.16: Human capital z (White)

Concomitant with declines in $\hat{\eta}$, Black investment increases, with little change in White investment. The non-effect on White investment might come as a surprise, but the increase in competition has offsetting effects at the extensive and intensive margins. As competition increases, investing in a small amount of human capital is less profitable, because there is stiff competition. However, conditional on investing, White workers raise their investment to increase the likelihood that they are hired and the resulting pay should they gain employment. These trade-offs will be important when we examine Affirmative Action.

Last, we show that firms update their beliefs and reduce the bias in their priors over time. To do this, we plot two figures in Figure 5.1.2. On the left hand side, we plot the

change in the average human capital of a firm's Black applicants against the change in the bias (E[z] - z) of that firm. Because firms initially underestimate Black human capital, the positive relationship this figure shows indicates that as firms encounter more well-qualified Black applicants, they reduce their bias and raise their expected human capital. In the right panel, we plot the correlation between an applicants true human capital, z, and the firm's expected human capital, $E[z]$. Because the signal received by a firm is composed of the true value and iid measurement error, $E[z]$ and z should be perfectly correlated; however, biased beliefs yield a correlation of only around 70% initially. Over time, beliefs update and this correlation rises.

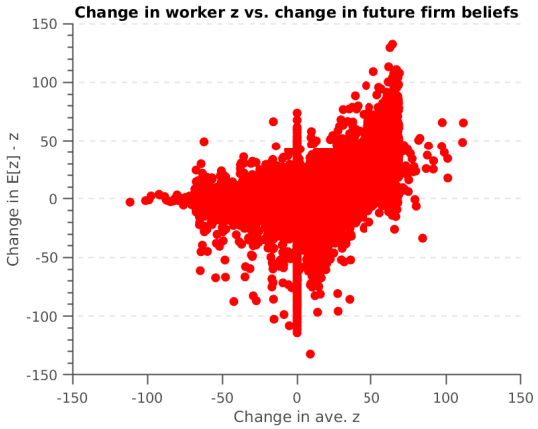


Figure 5.17: Change in firm z applicants vs. next period change in E[z] - z.

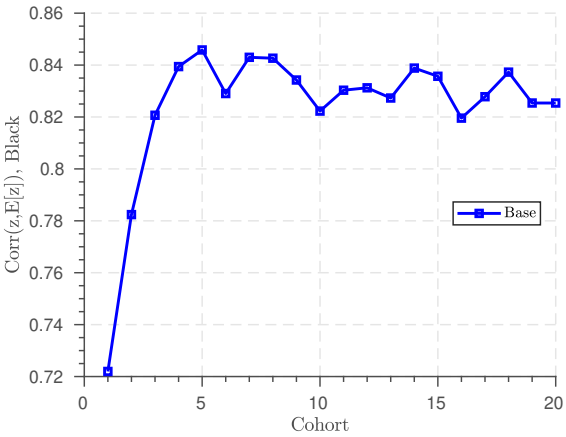


Figure 5.18: Correlation between E[z] and z

5.2 The Effects of Affirmative Action

In this section, we answer the question posed at the beginning of the paper: can temporary Affirmative Action employment policies have long-term benefits by affecting worker and firm beliefs? To assess this question, we impose a realistic Affirmative Action policy that affects a single cohort. Any effects on subsequent cohorts occur through changes in the information garnered from interviewing and hiring more Black candidates by firms, and by changes in the overall hiring makeup of the economy.

In our counterfactual, our Affirmative Action policy raises the score of all Black applicants by a fixed amount, similar to the policies instituted by many public universities prior to the Grutter v. Bollinger (2003).[8] Without an explicit target, we choose to set the hiring incentive, ζ, to be equal to the penalty applied to Black workers by taste-based discriminators, $\zeta = \gamma$. It is important to note that we impose this hiring incentive during the first cohort after investment decisions have been made, which means that Black workers are unaware of the favorable labor market when choosing their human capital.

We focus on the impact of Affirmative Action over two horizons: the immediate, short-run effects, and the persistent long-run effects. The reason for this dichotomy is twofold: first, the short-run effects depend on initial conditions, while the long-run effects depend on the manner and extent to which beliefs update. Second, the discussion of Affirmative Action often centers on short-run effects, while the cumulative effect over the long-run is likely to dominate any short-run outcomes. We proceed chronologically, beginning by examining the short-run effects in our calibrated model and then assessing the long-run effects, before concluding with the overall effects.

5.2.1 Short-Run Effects

In our model, Affirmative Action has two possible effects over a short horizon, depending on the initial state of the economy. First, firms may receive applications from woefully underqualified Black candidates and be compelled to hire them despite interviewing more qualified White candidates. In such an environment, firms may revise down their beliefs about Black productivity and negatively affect future generations of Black workers. This widely-held belief (see Sowell (2004) among others) is perhaps the most common argument against augmenting the current slate of Affirmative Action policies. By contrast, a second possible outcome allows firms revise up their beliefs about Black productivity. This occurs

[8]While an explicit "points-based" system may not pass the strict scrutiny test applied to Affirmative Action laws by courts, our approach is to abstract from more holistic considerations, which would be consistent with the precedent set down by Fisher v. University of Texas at Austin (2013).

if the history of hiring is so limited that firms interpret strong signals of Black productivity to be composed primarily of noise. By inducing additional Black hires, firms revise up their beliefs, where they otherwise would have had little scope to do so.[9]

We find evidence that more closely aligns with the first narrative about Affirmative Action, although it is important to note that these effects may not be present if we assumed that the Affirmative Action policy were announced in advance of investment decisions. In the short-run, Affirmative Action has two direct, tangible costs that occur because many White workers are displaced by less qualified Black workers. We present the impact of Affirmative Action on hiring and firm beliefs for the first cohort in Figure 5.2.1. In the left panel, we present a scatter plot of the human capital of displaced White workers, and the human capital of Black workers hired in their stead. In the right panel, we present the impact on the average firm posterior over Black productivity.

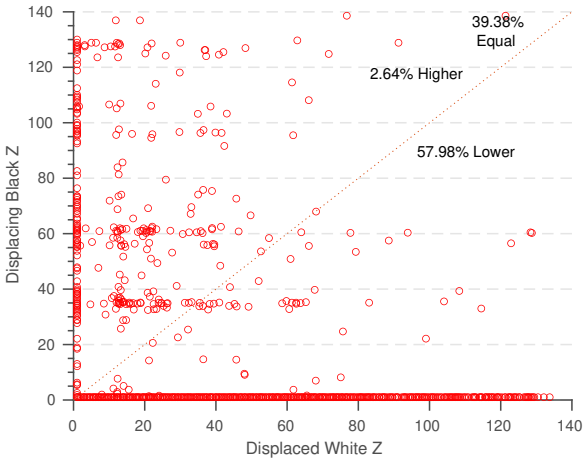


Figure 5.19: Human capital of Black hired and White displaced workers.

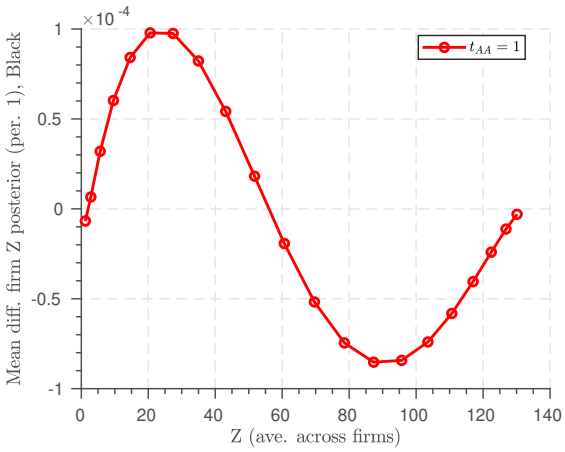


Figure 5.20: Difference in firm beliefs.

Figure 5.21: Short-run impact on hires and beliefs.

Both figures deserve careful analysis. In the left panel, we focus exclusively on hires in our Affirmative Action experiment in which a Black worker was hired by a firm that hired a White worker in our baseline economy absent the Affirmative Action benefit. While some

[9]While this second outcome is possible, the structure of the model limits its likelihood. It would require Black workers to be optimistic enough to invest in human capital, but still be rarely hired.

new Black hires are revealed to have higher human capital, nearly 60% of the displacing hires result in a Black worker with less human capital than the White worker they displaced. This occurs because a large share of Black workers invest very little in human capital prior to the institution of our Affirmative Action policy, but apply and are hired as a result of the policy. It is still worth noting that our results lend some credence to the beliefs about statistical discrimination: 2.64% of matches result in a Black candidate with more human capital than the White worker being displaced, and roughly 40% end with no change in human capital.

In the right panel, we calculate the difference in the average posterior of firms in our Affirmative Action economy and in our baseline economy after interviewing and hiring workers from the first cohort. Firms place more weight on low values of human capital for Black workers, that they rarely encountered in the baseline economy, and less weight on higher values of z. This indicates that Affirmative Action indeed exacerbates the statistical discrimination present in the labor market in our calibrated economy.

Much of the discourse, about Affirmative Action, both academic and public, ends upon reaching this conclusion. This static analysis, while valuable, ignores the prospect that the resulting equilibrium can yield tangible changes in the economy if beliefs do not fully align on both sides of the market. We extend this analysis in the next section by assessing the changes in beliefs and resulting changes in outcomes over time.

5.2.2 Long-Run Effects

Can Affirmative Action yield long-term benefits to racial inequality and the aggregate economy? Do short-term negatives preclude long-term positives? Our environment provides a unique laboratory in which to consider this question. In the previous section, we showed that for the treated cohorts, Affirmative Action yields economically negative outcomes. As we show in this section, extending this analysis to include subsequent cohorts is crucial to fully quantify the ramifications of Affirmative Action.

We begin by examining the impact that the labor market outcomes in the previous

section have on Black beliefs about discrimination and employment. In Figure 5.2.2, we plot the impact that drastic changes in hiring have on discrimination beliefs, $\hat{\eta}$, and as a result, on beliefs about employment prospects. In the left panel, we show a scatter plot of the percentage point difference in hiring between the baseline and the Affirmative Action economy along with the resulting difference in $\hat{\eta}$ on the vertical axis. In the right panel, we plot the second cohort's beliefs about employment probabilities for the baseline economy (blue line) and Affirmative Action economy (red line).

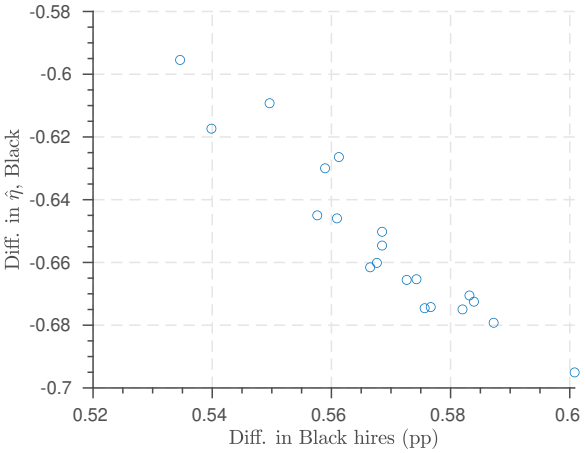


Figure 5.22: Diff. in $\hat{\eta}$ from Base across different runs.

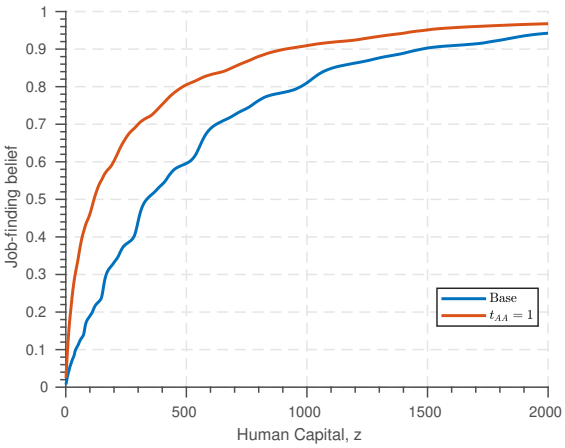


Figure 5.23: Diff. in $P(e|B, \hat{\eta})$.

Both figures show the sizable effect of Affirmative Action on beliefs. Despite the knowledge that the Affirmative Action policy would cause a sizable shift in employment, Black workers underpredicted the extent to which an Affirmative Action policy would increase Black employment. As a consequence, discrimination beliefs fall by between 60 and 70 percentage points. This change in beliefs about discrimination causes the current generation to be substantially more optimistic about their employment prospects: in the baseline, a Black worker with approximately average investment ($z = 20$) believed they had a 5% chance of receiving a job offer. Now, after revising $\hat{\eta}$ in response to the employment shift induced by Affirmative Action, Black workers believe that if they pay the cost to acquire this human capital, they have a more than 20% chance of obtaining high prestige employment.

As shown in Figure 5.1.1, beliefs about employment probability play a large role in human capital investment decisions. How does the sizable shift in employment beliefs affect human capital investment here? The left panel in Figure 5.2.2 compares the investment decisions for the second cohort of Black workers in the Affirmative Action economy (red circles) to the second cohort of Black workers in the baseline economy (blue circles). The right panel compares Black and White investment decisions for the second cohorts in the Affirmative Action economy.

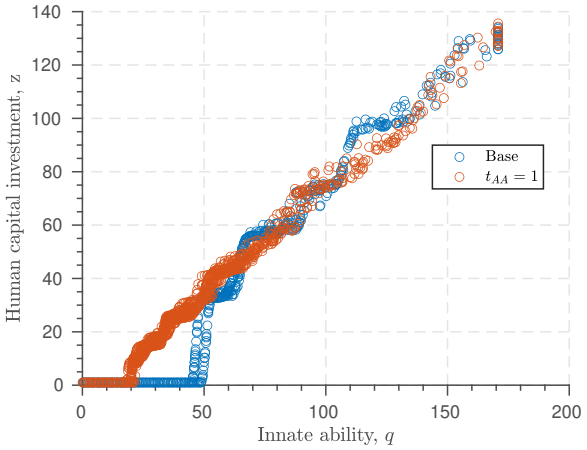


Figure 5.24: Human capital investment decisions (z) by innate ability (q) (Black).

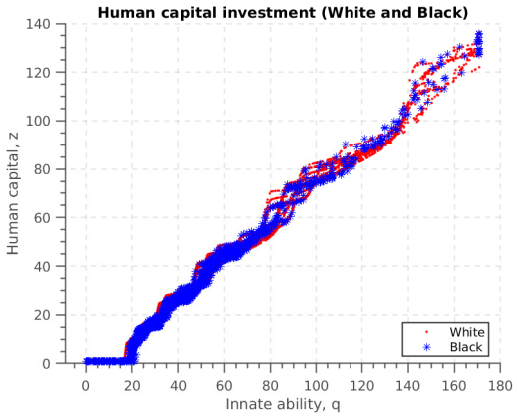


Figure 5.25: Human capital investment decisions (z) by innate ability (q) and race.

Although they are no longer propped up by a score boost, Black workers increase their investment in human capital nearly across the innate ability distribution. The left panel shows that many more Black workers are willing to invest in human capital than in the Baseline economy. Perhaps more surprisingly, the right panel shows that Black human capital investment now closely mirrors White investment, while one can easily infer from the Baseline investment in the left panel that this is not the case in the absence of the Affirmative Action policy.

Yet, these gains may only be transitory. As Section 5.2.1 showed, revised their beliefs about Black productivity down, meaning that the increase in investment may fall on the deaf ears of hiring managers as they filter signals from Black candidates. But this turns out

not to be the case. In Figure 5.2.2 we plot two figures that show the resulting changes in firm beliefs following the interviewing and hiring of workers from the second cohort. In the left panel, we plot the difference in firm posteriors between the Affirmative Action economy and the Baseline economy over Black productivity. In the right panel, we plot (z, $E[z]$) pairs for the Affirmative Action (red circles) and Baseline (blue circles) economies.

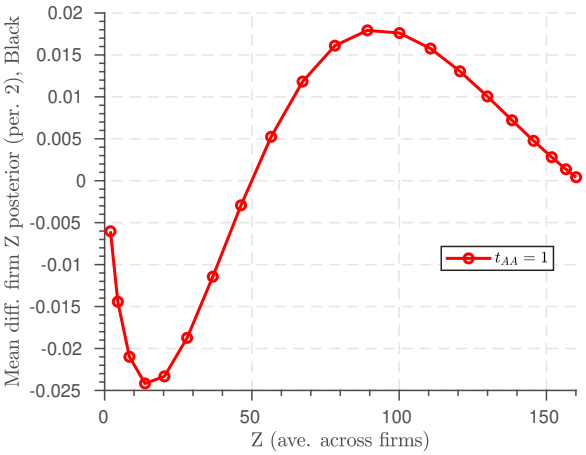


Figure 5.26: Diff. in firm z beliefs (Black), gen. 2

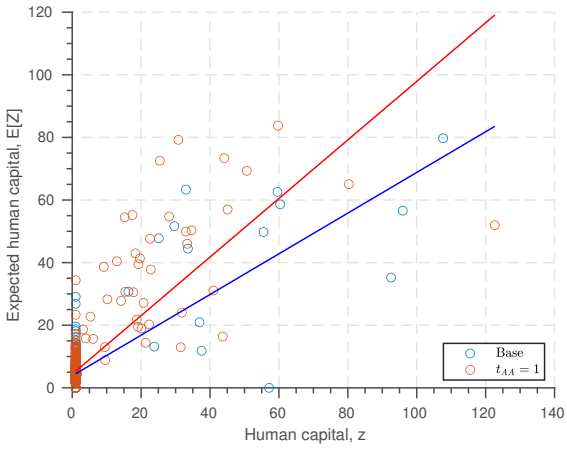


Figure 5.27: Average z

Despite backtracking during the first cohort, Black workers make sizable gains across the human capital distribution. The left panel shows that firms place more weight on strong signals from Black applicants indicating high human capital rather than noise. The right panel shows that bias in beliefs declines, as the line of best fit in the Affirmative Action economy nearly perfectly matches the 45 degree line.

It is easy to infer from the near immediate improvements in outcomes for Black workers that these gains persist. And indeed, they do. In Figure 5.29, we plot $\hat{\eta}$ in the left panel and hiring probabilities in the right panel for the Baseline (blue) and Affirmative Action (red) economies.

The left panel shows that $\hat{\eta}$ experiences an immediate and sizable decrease in the Affirmative Action economy. It even drops below the measure of taste-based discriminating firms ($\mu = 0.223$), indicating that Black workers outperformed their own predictions. This occurs

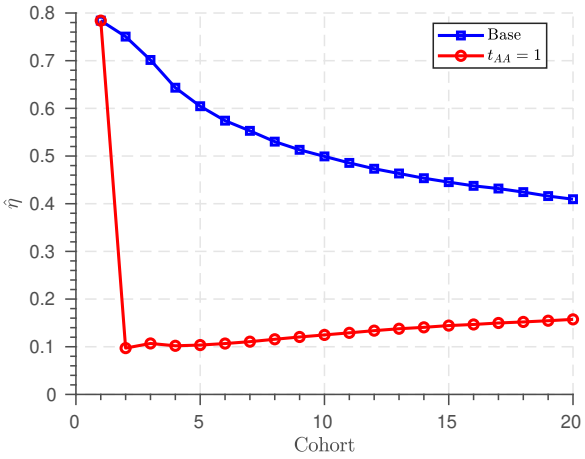


Figure 5.28: $\hat{\eta}$ across cohorts

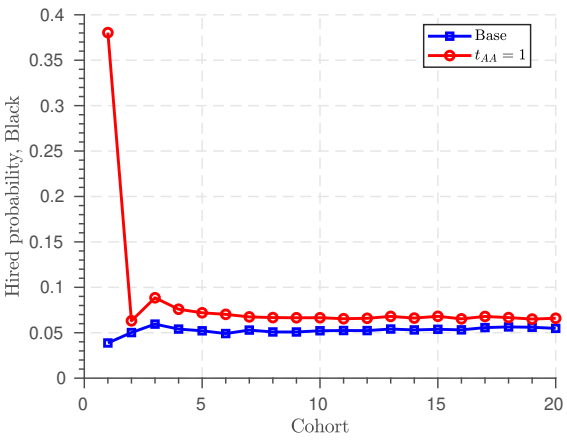


Figure 5.29: Black emp. outcomes

because many applicants invested very little in human capital and prior to the Affirmative Action policy would not have applied; now, they are hired and firms assign positive weight to higher levels of human capital, meaning that statistical discrimination is benefitting these workers. As a result of the reduction in discrimination beliefs and the increase in human capital investment, the right panel shows a persistent increasing in hiring probability for Black workers.

While these gains are undoubtedly important, such a policy may be less appealing in positive terms if the gains experienced by Black workers come at the expense of White workers. In Figure 5.2.2, we plot average Black and White human capital and compare to the Baseline in each figure.

The left panel shows a clear and persistent increase in human capital for Black workers. Surprisingly, the right panel appears to exhibit at least a small positive effect on White human capital as well for most cohorts. The reason is that there are two implicit dimensions to the investment decision: first, an extensive decision about whether or not to invest at all, and second, an intensive decision about how much to invest. In this case, increased competition causes White workers to invest more in their human capital, which dominates the extensive effect on any human capital investment.

Last, we present the average effect on the economy for several key variables in Table 5.1.

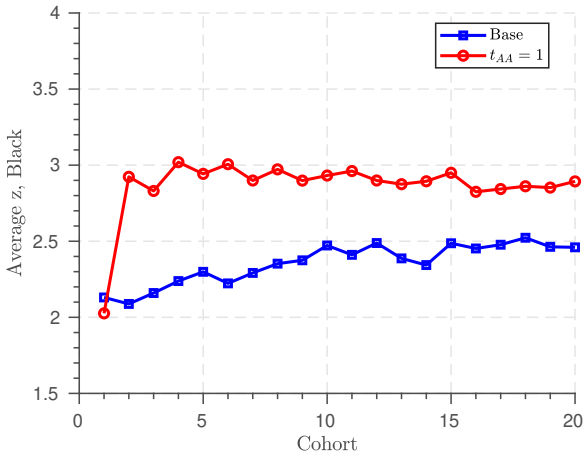


Figure 5.30: Human capital z (Black)

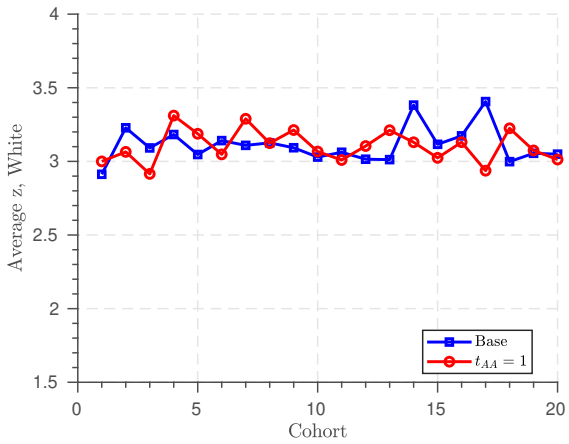


Figure 5.31: Human capital z (White)

Value	Base	$t_{AA} = 1$
Ave. $\hat{\eta}$	0.5240	0.1610
Employed Z (Post AA)	18.627	18.913
Black Z Invest	2.359	2.865
White Z Invest	3.162	3.139
Black Employed Z (Post AA)	19.565	19.646
White Employed Z (Post AA)	18.536	18.819
Black High Prestige Emp. Rate	0.0530	0.0840
White High Prestige Emp. Rate	0.0910	0.0860
Average Black Income	1.857	2.127
Average White Income	2.480	2.387

Table 5.1: Average aggregate effects.

Affirmative Action proves to be highly effective at improving outcomes for Black workers, while imposing only slight costs on White workers. As a result of the Affirmative Action policy, the racial income gap falls by 14 percentage points, or nearly 57%. Black high prestige employment is on par with White employment in the same jobs and the human capital of both Black and White employed workers increases. Investment increases by 21.5% for Black workers, while the decrease for White workers is less than 1 percent (0.7%).

5.3 The Importance of Two-Sided Beliefs

In this section we demonstrate that the inclusion of two-sided updating beliefs is vital for obtaining our quantitative findings in Section 5.2.1 and Section 5.2.2. We consider 3 restrictions. First, we restrict workers to believe that discrimination always remains at its initial level, $\hat{\eta}_t = \hat{\eta}_0 \forall\ t$, where t refers to the cohort. Second, we restrict the beliefs of firms in our economy to remain at their initial distributions, $f(z|y,r)_{i,t} = f(z|y,r)_{i,0} \forall\ i,\ t$, where i refers to the firm. Last, we impose both restrictions to understand the importance of the interaction between beliefs. We consider the first two counterfactuals in Section 5.3.1 and then assess the importance of their interaction in Section 5.3.2

5.3.1 Firm and Worker Beliefs

We first examine the role of firms updating their beliefs in our findings. In our baseline economy, firm updating created an echo effect, amplifying the decisions of workers by updating their beliefs positively when workers invest more in human capital and negatively when workers invest less. Here, however, their initial degree of statistical discrimination is permanent, mitigating this amplification. Does it materially alter the equilibrium? We start by plotting the time series of discrimination beliefs and the probability a Black worker receives a job offer, relative to our Baseline economy in Figure 5.3.1. In both figures the Baseline economy is denoted by a blue line with square markers and the $f(z|y,r)_{i,t} = f(z|y,r)_{i,0}$ is denoted by a red line with circular markers.

While $\hat{\eta}$ declines in both economies, it declines much more rapidly in the baseline, demonstrating the role that statistical discrimination plays in causing belief retrenchment of Black cohorts. On the right, the hiring probability figure demonstrates the first departure from the Baseline: while the fixed beliefs economy exhibits slight upward movement in hiring probability, this is due to changes in Black investment.

Does this affect Black investment in human capital? This depends on the degree of

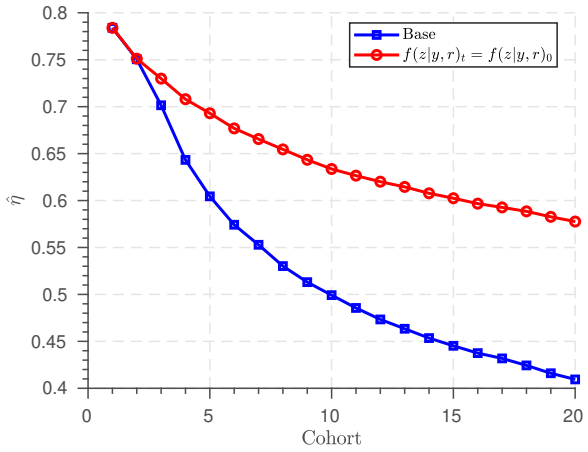


Figure 5.32: $\hat{\eta}$.

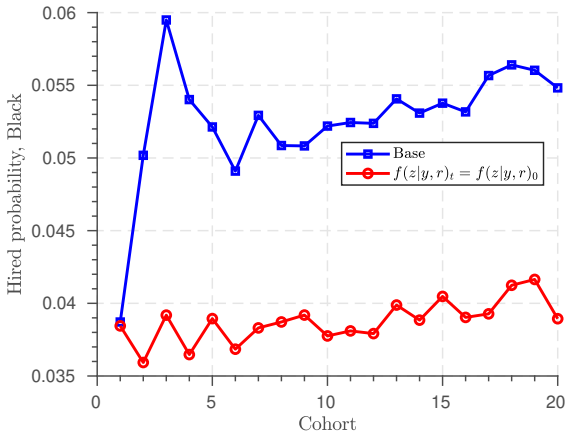


Figure 5.33: Employment probability (Black).

statistical discrimination, which Figure 5.32 suggests a is large. We plot these series in Figure 5.3.1, with Black human capital on the left and expected Black human capital of those who apply averaged across firms on the right.

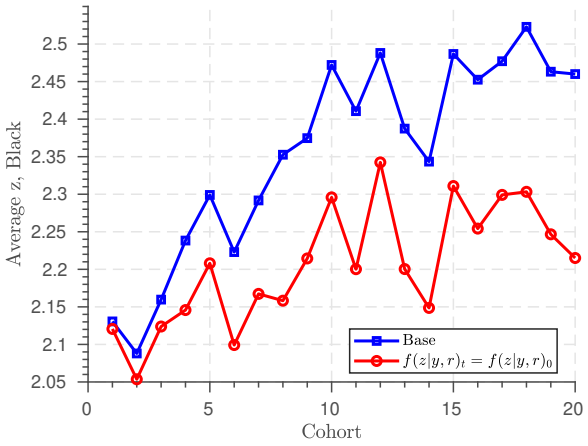


Figure 5.34: Average z.

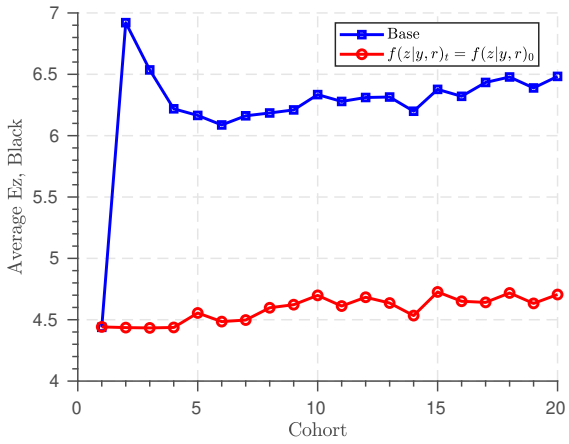


Figure 5.35: Firm $E[z]$ of Black applicants.

While Black investment increases, it does so at a slower pace than in the Baseline economy. This, however, translates into almost no change in firm expectations over Black human capital, yielding no reward for increased investment by Black workers.

Now we consider a counterfactual in which worker beliefs are fixed at their initial value, $\hat{\eta}_t = \hat{\eta}_0 \forall\ t$, while firm beliefs are allowed to vary. The impact in this context is much more evident: being unable to update their beliefs, Black workers simply never adjust to

the dynamics of the labor market and languish behind their accomplishments in the baseline economy. We first plot Black employment probabilities along with the average level of human capital in Figure 5.3.1.

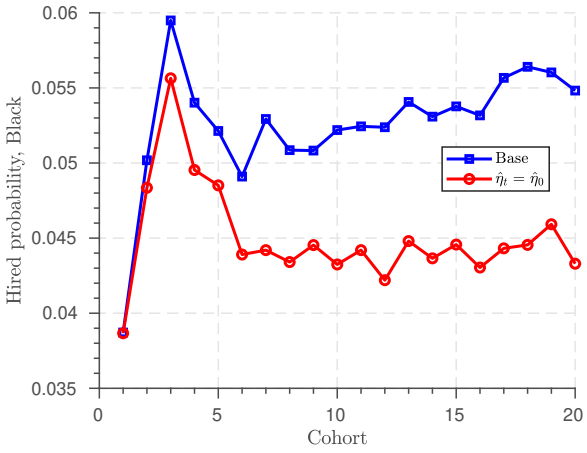


Figure 5.36: Employment probability (Black).

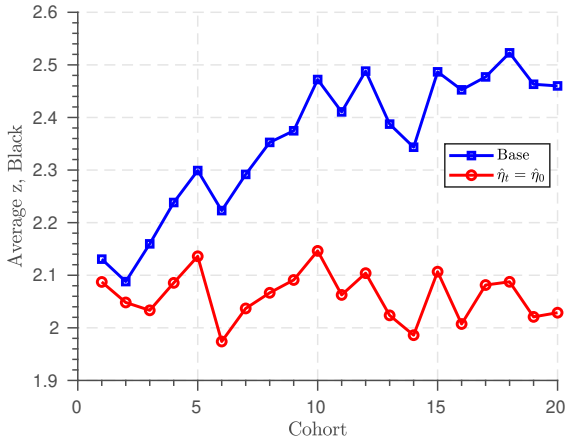


Figure 5.37: Average z.

While there is an initial increase in hiring probability as firms adjust their beliefs, it rapidly declines and stagnates just above its initial value. The time series for z in the counterfactual looks even less promising: with limited exceptions, each subsequent cohort averages lower human capital than their predecessors.

To highlight just how tangible this restriction becomes, we plot the beliefs over employment prospects (left) as well as the difference in average firm beliefs (right), both for the 10th cohort, in Figure 5.3.1. In the left figure, the blue line corresponds to the baseline economy and the red line corresponds to the restricted economy.

Although hiring decisions depend principally on firm beliefs, given human capital, Black workers are universally more pessimistic. The figure on the right shows the effect of persistent underinvestment: firms upweight the very lowest levels of human capital, while downweighting higher levels. In this figure, it's worth highlighting that the peak difference occurs below the average level of human capital investment, which occurs around $z = 20$. This indicates that firms are pessimistic that Black applicants even achieve this level of human capital.

How do these restrictions change the effectiveness of Affirmative Action? We repeat our

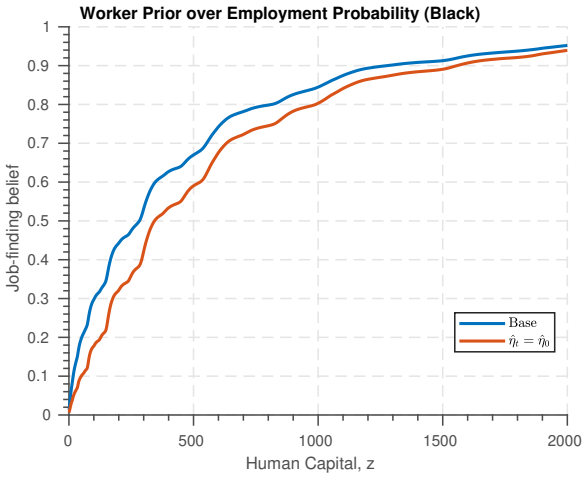


Figure 5.38: Black employment beliefs (cohort 10).

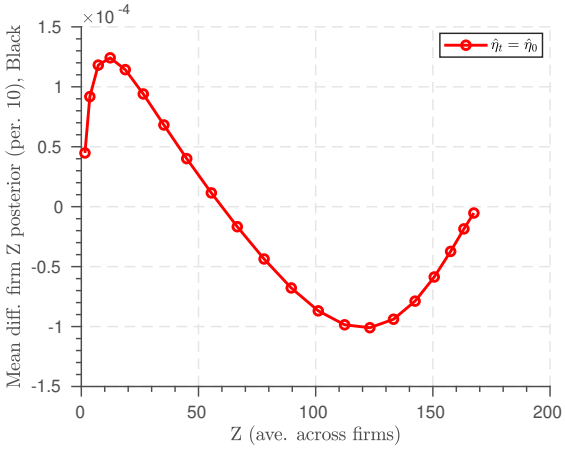
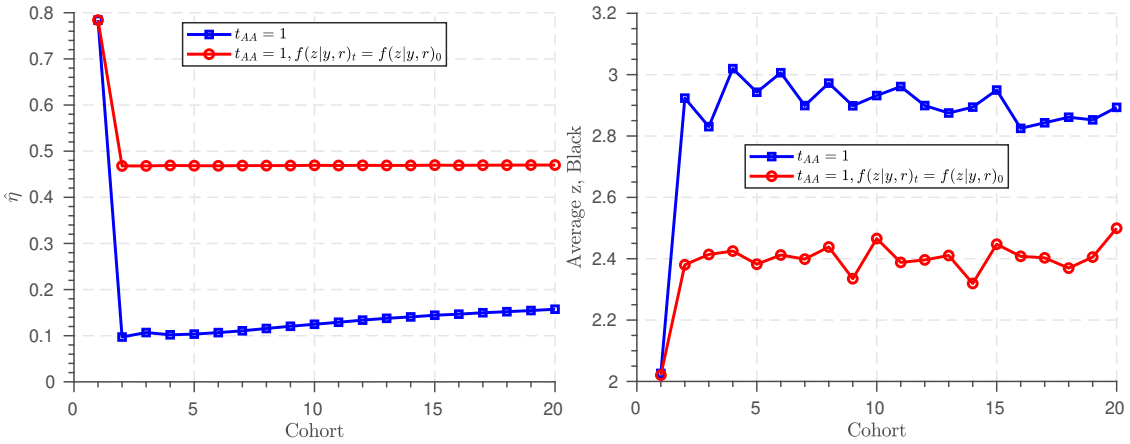


Figure 5.39: Difference in average firm beliefs (cohort 10).

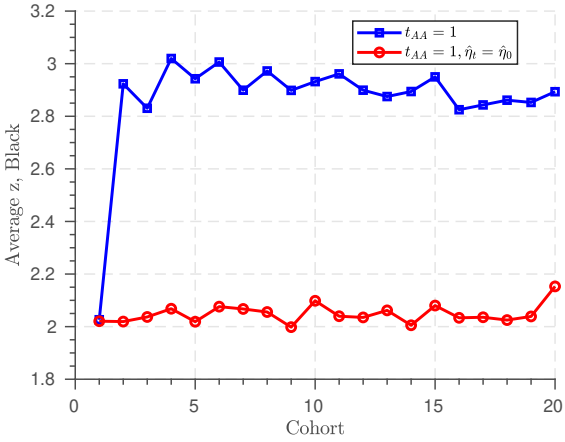
experiment from Section 5.2 under both restrictions and plot the results for key quantities in Figure 5.40. In the upper panels, we plot $\hat{\eta}$ (left) and average Black human capital, z, (right) under the counterfactual in which firm beliefs are restricted. In the lower two panels, we plot average Black human capital, z, (left) and Black beliefs over employment prospects at the start of the second cohort (right) under the counterfactual in which we restrict worker beliefs.

Each figure provides stark evidence of the important role played by updating beliefs. In the firm belief counterfactual, Affirmative Action has an immediate impact on $\hat{\eta}$; however, it only falls to a value slightly below 0.5 and remains there for the rest of the simulation. This is because statistical discrimination is permamently embedded in the economy and remands much of the progress accomplished by Affirmative Action in the baseline economy. The right top panel provides further evidence to this point: while Affirmative Action causes increased Black investment, the inability of firms to respond to this change cuts the benefits roughly in half. The bottom two panels show an even more seismic effect: when workers cannot update their beliefs, the left panel shows that Affirmative Action has no effect on Black human capital investment. The right panel shows that despite unrestricted beliefs, firms hold persistently more pessimistic beliefs as a result of the lack of Black human capital

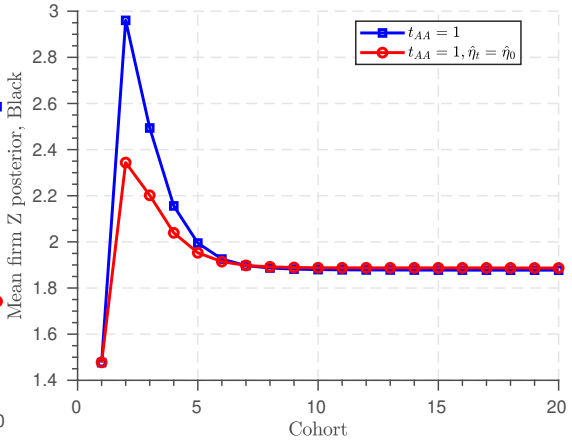


(a) $\hat{\eta}$ under restricted firm beliefs.

(b) Black average z under restricted firm beliefs.



(c) Black average z under restricted worker beliefs.



(d) Mean firm beliefs over Black human capital under restricted worker beliefs.

Figure 5.40: Affirmative Action outcomes under restricted beliefs.

investment than they do when both beliefs are unrestricted.

The upper left and bottom right panel allude to the impact that restricting beliefs on one side of the market can have on the other. $\hat{\eta}$ never achieves the same levels when firm beliefs are restricted, and when worker beliefs are restricted, the lack of investment makes firms more pessimistic. In the next section, we explore the importance of the interaction between firm and worker beliefs.

5.3.2 Interaction

Now we assess the interplay between beliefs. In Section 5.1 we explored the role of feedback between worker optimism, their human capital investment, and the subsequently firm posteriors over worker productivity. Now, we formally quantify the importance of this feedback loop. To do this, we consider a third restriction in addition to our two restrictions from the previous section. In this new restrictions, beliefs on both sides are restricted to their initial values. This provides a fully restricted counterfactual against which to compare the other restrictions and our unrestricted baseline model. We refer to this new restriction as "R1," the counterfactual in which firm beliefs are restricted as "R2," and the counterfactual in which worker beliefs are restricted as "R3." We use "UR" to denote our unrestricted baseline model.

The difference between these restrictions allow us to quantify the importance of each updating channel. The difference in outcomes between UR and R1 provides the cumulative effect of allowing beliefs to update. Because R2 holds firm beliefs fixed, taking the difference between R1 and R2 yields the effect of allowing worker beliefs to update in isolation. The same logic applies to the difference between R3 and R1, which yields the role of updating firm beliefs.

Any interaction that occurs between updating beliefs occurs through indirect channels. When workers become more optimistic, they raise their investment; this alters the views held by firms about their productivity. These updated views cause firms to hire more workers, which increases optimism, and so on. Our restricted counterfactuals limit the effect of beliefs to the direct effect: in these contexts, only the next link in this feedback chain occurs. This allows us to quantify role of the interaction by taking the difference between the cumulative effect (UR - R1) and the sum of the firm (R3 - R1) and worker (R2 - R1) effects. We present these results as well as the percent of the cumulative effect explained by the interaction in Table 5.2.

Value	R2-R1	R3-R1	UR-R1	Interaction	Pct. Explained
Ave. $\hat{\eta}$	-0.1400	0.0000	-0.2603	-0.1203	46.216
Employed Z	0.0660	0.2820	0.4280	0.0800	18.735
Black Z Invest	0.1430	0.0009	0.2940	0.1500	51.185
White Z Invest	-0.0053	-0.0027	-0.0224	-0.0143	63.850
Black Employed Z	0.6450	0.6930	1.283	-0.0555	-4.3215
White Employed Z	0.0260	0.2470	0.3430	0.0700	20.498
Black High Prestige Emp. Rate	0.0022	0.0085	0.0160	0.0053	33.136
White High Prestige Emp. Rate	-0.0004	-0.0014	-0.0027	-0.0009	32.881
Average Black Income	0.0640	0.1800	0.3480	0.1040	29.928
Average White Income	-0.0040	-0.0020	-0.0152	-0.0092	60.460

Table 5.2: Decomposition of model mechanism into updating worker and firm beliefs, and their interaction.

This table shows that the interaction between updating beliefs plays an important role for nearly all of our key variables. The most direct measure of its role is on discrimination beliefs: on their own, changes in worker beliefs account for a 14 percentage point decrease in $\hat{\eta}$; allowing firms to also update their belief yields a further 12 percentage point decrease, indicating the interaction accounts for nearly 50% of the overall effect. Among the remaining variables, the interaction explains similar magnitudes except for employed human capital. Notably, for Black workers, the interaction reduces the gains accrued from updating beliefs separately by 4.32%. The reason is subtle: when few Black workers invested in human capital, firms rarely received competitive applications from any Black workers except those with high innate ability and corresponding high human capital. Allowing beliefs to update cause Black workers with lower human capital to apply and firms to place more weight on lower levels of productivity.

5.4 Can Affirmative Action Backfire?

In this section, we consider an obvious corrolary to our findings: Are there circumstances in which Affirmative Action can have a negative impact on long-term Black outcomes? While our calibration indicates that Affirmative Action provides extensive and permanent benefits to Black workers and the economy as a whole, we indeed find that there are parametrizations for which Affirmative Action backfires. When Affirmative Action reveals that the labor

market is worse for Black workers, it can cause stark upward revisions of beliefs and a long-term decline in human capital investment. In this section, we explore two specific calibrations and give an overview of calibrations in which Affirmative Action is no longer effective.

Perhaps the most obvious circumstance in which Affirmative Action backfires is when there are is a large measure of taste-based discriminating firms in the economy. In this environment, Affirmative Action reveals to Black workers what would become evident over a much longer horizon: that many firms are virtually unwilling to hire Black workers, rendering the return on human capital investment negligible. For our experiment, we set the measure of taste-based discriminators to $\mu = \hat{\eta}_0 = 0.784$, the circumstance in which initial beliefs about discrimination reflect entirely taste-based discrimination. We plot the resulting discrimination beliefs (left panel) and human capital investment of Black workers (right panel) in Figure 5.43. In both figures, the baseline economy (with the new value of μ) is plotted in blue, while the Affirmative Action embedded economy is in red.

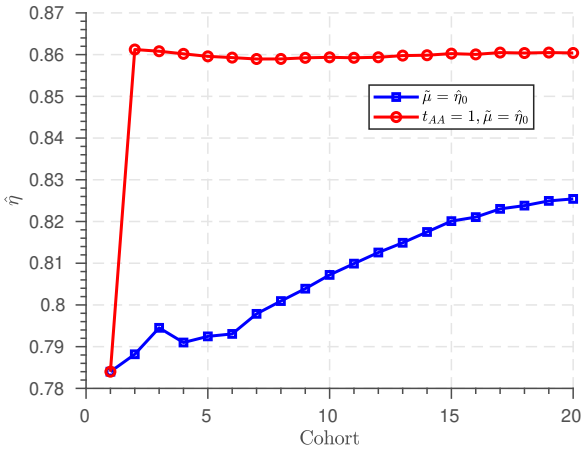


Figure 5.41: $\hat{\eta}$.

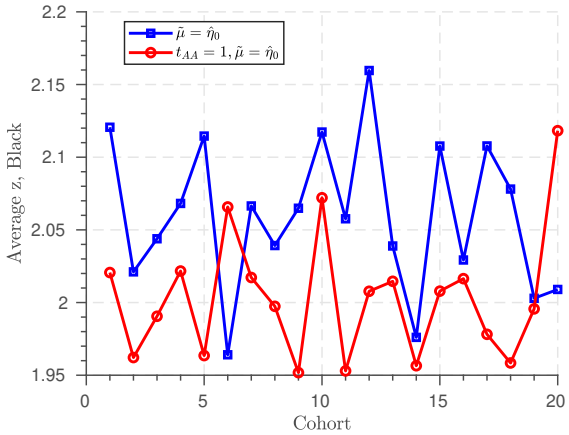


Figure 5.42: Black human capital, z

Figure 5.43: The effect of Affirmative Action under $\tilde{\mu} = \hat{\eta}_0 = 0.784$

Next, we adjust worker beliefs so that they are accurate about the measure of taste-based discriminators in the economy, so that $\hat{\eta}_0 = \mu = 0.223$. While dynamically, this yields the unique ergodic equilibrium, the two-sided nature of our environment prevents this outcome from arriving immediately. Because firms statistically discriminate, believing that

discrimination only arises from taste-based firms understates the degree of discrimination in the labor market, leading to negative revisions of beliefs. We plot the resulting discrimination beliefs (left panel) and human capital investment of Black workers (right panel) in Figure 5.46. As before the baseline economy (with the new value of μ) is plotted in blue, while the Affirmative Action embedded economy is in red in both figures.

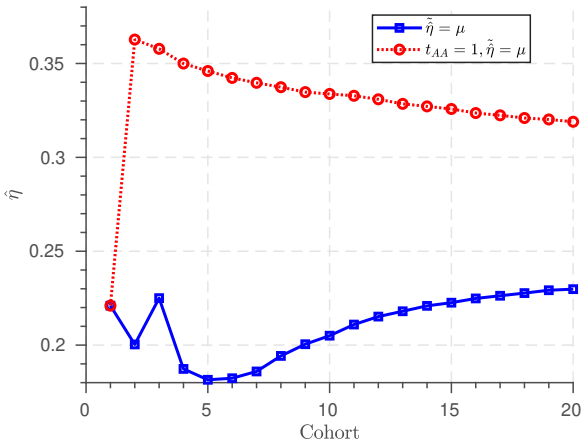


Figure 5.44: $\hat{\eta}$.

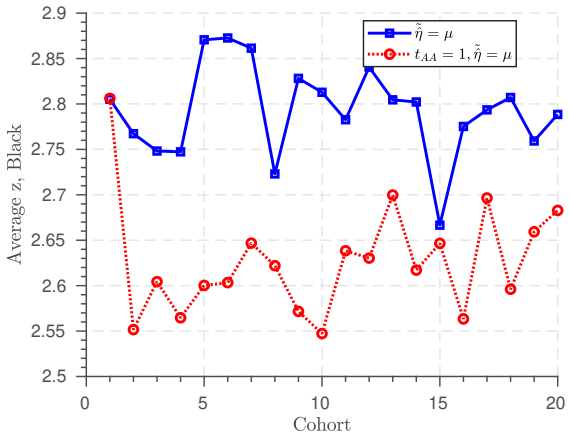


Figure 5.45: Black human capital, z

Figure 5.46: The effect of Affirmative Action under $\tilde{\hat{\eta}}_0 = \mu = 0.223$

Again, Affirmative Action negatively affects Black workers. A noteworthy subtlety is that the negative effect is larger in this second experiment; the reason is that in the first experiment, there is very little scope for policy interventions to have any effect. The measure of taste-based discriminators is so large that Black workers are already extremely pessimistic, whereas they are relatively optimistic in the second experiment. This leaves more room for statistical discrimination to have an impact.

In both cases, the reason that Affirmative Action backfired is that Black workers implicitly attributed some share of labor market discrimination to statistical discrimination. In each of these environments, taste-based discrimination already accounts for as much or more discrimination than Black workers infer from the labor market; the consequence is that Affirmative Action reveals that the labor market is more hostile than they previously believed. This is a general conclusion: when Black workers do not account for statistical discrimination,

Affirmative Action reveals additional racial animus in the labor market. The result is that in these environments, Affirmative Action is not effective.

6 Conclusion

Racial economic inequality is an issue of paramount concern. In this paper, we show that Affirmative Action provides an effective avenue for addressing it, even when only enacted over a short horizon. Despite negative short-term effects, Affirmative Action provides substantial long-term benefits.

We reach this conclusion by constructing a model in which workers and firms face information asymmetries: Black workers believe a large share of firms are taste-based discriminators which means that the majority forgo accumulating human capital, believing that this investment would prove fruitless. As a consequence, even non-discriminatory firms place little faith in positive signals from Black applicants. The resulting absence of Black hires leaves firms with little opportunity to revise their beliefs, and reinforces Black workers' beliefs about the extent of discriminatory they face in the labor market. In this environment, Affirmative Action can play a crucial role: it decouples the link between biased firm priors and employment outcomes, which yields information to Black workers about the degree of "true" discrimination in the labor market.

Like much of the previous literature, we find that after an Affirmative Action is instituted Black workers displace more qualified White workers and firms revise their priors down. But this only occurs in the short-run. Unlike the previous literature, we consider the dynamic effects of Affirmative Action on the beliefs of future generations. By observing the labor market success of the previous generation, future cohorts learn that there are fewer taste-based discriminators than believed and that the remaining firms are willing to hire Black workers who appear qualified. This results in more investment by future generations of Black workers, which becomes self-reinforcing as firms learn that the average Black applicant

is more capable than they previously believed. Ultimately, these long-term benefits far outweigh the short-term costs and lead to higher productivity for both Black and White workers as well as the economy as a whole.

References

Arcidiacono, Peter, Esteban M Aucejo, Hanming Fang, and Kenneth I Spenner, "Does affirmative action lead to mismatch? A new test and evidence," Quantitative Economics, 2011, 2 (3), 303–333.

Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor, "Does affirmative action work? Caste, gender, college quality, and academic success in India," American Economic Review, 2016, 106 (6), 1495–1521.

Becker, Gary S, The economics of discrimination, University of Chicago press, 1957.

Blume, Lawrence E, "The dynamics of statistical discrimination," The Economic Journal, 2006, 116 (515), F480–F498.

Boerma, Job and Loukas Karabarbounis, "Reparations and persistent racial wealth gaps," Technical Report, National Bureau of Economic Research 2021.

Chung, Kim-Sau, "Role models and arguments for affirmative action," American Economic Review, 2000, 90 (3), 640–648.

Coate, Stephen and Glenn C Loury, "Will affirmative-action policies eliminate negative stereotypes?," The American Economic Review, 1993, pp. 1220–1240.

Duncan, Otis Dudley, "A socioeconomic index for all occupations," Class: Critical Concepts, 1961, 1, 388–426.

Fang, Hanming and Andrea Moro, "Theories of statistical discrimination and affirmative action: A survey," Handbook of social economics, 2011, 1, 133–200.

_ and Peter Norman, "Government-Mandated Discriminatory Policies: Theory and Evidence," International Economic Review, 2006, 47 (2), 361–389.

Fischer, Mary J and Douglas S Massey, "The effects of affirmative action in higher education," Social Science Research, 2007, 36 (2), 531–549.

Fisher v. University of Texas at Austin

Fisher v. University of Texas at Austin, 2013.

Grutter v. Bollinger

Grutter v. Bollinger, 2003.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The allocation of talent and us economic growth," Econometrica, 2019, 87 (5), 1439–1474.

Levin, Jonathan et al., "The dynamics of collective reputation," The BE Journal of Theoretical Economics, 2009, 9 (1), 1–25.

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

Moro, Andrea and Peter Norman, "A general equilibrium model of statistical discrimination," Journal of economic theory, 2004, 114 (1), 1–30.

Phelps, Edmund S, "The statistical theory of racism and sexism," The american economic review, 1972, 62 (4), 659–661.

Sander, Richard H, "A systemic analysis of affirmative action in American law schools," Stan. L. Rev., 2004, 57, 367.

Shimer, Robert, "The cyclical behavior of equilibrium unemployment and vacancies," American economic review, 2005, pp. 25–49.

Sowell, Thomas, Affirmative action around the world: An empirical study, Yale University Press, 2004.

A Data Construction

A.1 National Longitudinal Survey of Youth, 1979 (NLSY79)

The National Longitudinal Survey of Youth follows cohorts who were ages 14-22 in 1979 through the present. It was conducted annually from 1979-1994 and bi-annually from 1994 until now, and includes detailed information on labor market status, including current employer, weeks employed, unemployed, and out of the labor force, as well as any training received by the individual since the last interview. Earnings are recorded annually as well as hours worked. In addition, the NLSY recorded a standardized test score, the Armed Forces Qualification Test (AFQT) for every individual in the sample. This allows me to link individuals by their AFQT scores to their outcomes late in the life-cycle.

A.2 Current Population Survey (CPS)

We use the Current Population Survey (CPS) to link labor market statistics to other outcomes of interest. We restrict the sample to be prime-aged males who are either Black or White. When constructing our moments for empirical analysis in Section 2, we use the CPS from 1972 to 1982, because this is the period in which it uses the 1970 occupation codes, like the NLSY79.

A.3 General Social Survey (GSS)

The General Social Survey is a broad social survey that has been conducted annually since the 1960s. It asks a wide variety of questions on common areas of interests to economists, employment, income, and includes all pertinent information on the demographics of its respondents. Importantly, it asks respondents questions about what they believe causes worse outcomes for Black workers in the labor market, categorized into "Discrimination," "Education," "Ability," and "Work Ethic." It asks a wide range of questions that allow the researcher to infer some information about the degree to which respondents have a taste-based preference for one race over another, like "Would you be comfortable having a person of the opposite race over for dinner?", "How would you feel about a family member marrying a person of the opposite race?" and many other related questions, each of which includes some variant of "Strongly Opposed" as a possible response.