Canadian Labour Economics Forum

WORKING PAPER SERIES

Endogenous Learning, Persistent Employer Biases, and Discrimination

Louis-Pierre LePage (University of Michigan)

CLEF Award 2021

WP #34

Endogenous Learning, Persistent Employer Biases, and Discrimination

Louis-Pierre Lepage*
University of Michigan

May 2021

Abstract

I present a new discrimination model of the labor market in which employers are initially uncertain about the productivity of worker groups and endogenously learn about it through their hiring. Previous hiring experiences of an employer shape their subsequent decisions to hire from a group again and learn more about its productivity, leading to differential learning across employers and biased beliefs about the group's productivity. Given a market-clearing wage, optimal hiring follows a cutoff rule in posterior beliefs: employers with sufficiently negative experiences with workers from a group stop hiring from the group, preserving negative biases and leading to a negatively-skewed distribution of beliefs about their productivity. When employers have noisier initial information on the productivity of one worker group, discrimination against that group can arise and persist without productivity differentials or prior employer biases, with market competition, and with or without worker signaling or investment decisions. The model generates steady state predictions analogous to the Becker (1957) taste-based model with beliefs replacing preferences, but is set within a statistical framework, explaining apparent prejudice as the result of "incorrect" statistical discrimination. The model also generates additional predictions and policy implications that contrast with previous models.

* lepagelp@umich.edu, Department of Economics, University of Michigan, Ann Arbor, 48109, US. I thank John Bound, Charlie Brown, David Miller and Jeff Smith for continued guidance and feedback. I thank Martha Bailey and Tanya Rosenblat for valuable advice and Dan Black, Aislinn Bohren, Kerwin Charles, Ash Craig, Ellora Derenoncourt, Pierre Deschamps, Alex Imas, Kevin Lang, Steve Lehrer, Erik Lindqvist, Conrad Miller, Nicolas Morales, Mike Mueller-Smith, Mandy Pallais, Rosina Rodriguez Olivera, Dan-Olof Rooth, Mel Stephens, Justin Sydnor, Joanne Tan, Justin Wolfers, Steve Woodbury, as well seminar participants at the CLEF/CEA, MEA, EALE/SOLE/AASLE, University of Michigan, Labo(u)r day at Michigan State University, University of Wisconsin-Madison, and EGSC at Washington University for their comments.

After decades of cultural change and anti-discrimination legislation, there remain substantial labor market outcome differentials across race and gender (Lang and Lehmann, 2012; Blau and Kahn, 2017). Two contrasting rationales are typically considered in economics to explain the contribution of discrimination to these disparities. Taste-based discrimination arises through exogenous preferences of employers for groups (Becker, 1957), creating differences between average performance and average pay of groups. Statistical discrimination instead arises as a rational response to true group productivity differentials when employers have imperfect information on individuals and therefore extract information from their group (Phelps, 1972; Arrow, 1973; Aigner and Cain, 1977, Coate and Loury, 1993; Moro and Norman, 2004). In these models, employers typically learn about individual productivity but not that of groups; they are assumed to know the productivity distribution of groups or at least to have correct equilibrium beliefs about it. In this paper, I propose an alternative explanation: discrimination can also arise from incorrect or biased employer beliefs about group productivity fueled specifically by their market interactions with groups.[1] Similarly to the taste-based literature, discrimination does not reflect true group differentials, but it also does not reflect a fundamental prejudice. Instead, discrimination arises from a lack of information or learning about groups, blurring the line between classical theories and leading to different predictions regarding how it arises and can be mitigated.

In particular, I posit that biased employer beliefs in the labor market arise naturally from initial uncertainty about the productivity of worker groups. That is, when employers enter the labor market, they are not only uncertain about the individual productivity of potential workers as in the statistical discrimination literature, but also the underlying productivity of their group. Since productivity may differ across groups, for example due to historical or social factors, employers value learning about groups to inform their hiring. A natural source of learning is their own hiring experiences with workers, implying that previous experiences of an employer with workers of a given group not only shape their beliefs about the group's productivity, but also their subsequent decisions to hire from the group and, indirectly, learn more about their productivity. Learning about minority or disadvantaged groups is particularly important if there is less initial information available about them in

[1] Evidence documenting biased beliefs in the context of gender, ethnicity, and race includes Fershtman and Gneezy (2001), Reuben et al. (2014), Mobius et al. (2016), Laouénan and Rathelot (2017), Van Dalen and Henkens (2019), Arnold et al. (2018), Landsman (2018), Lesner (2018), Bohren et al. (2019a, 2019b), Bordalo et al. (2019), Sarsons (2019) and Benson and Lepage (2021).

the labor market, making employers more reliant on their own experiences to assess their productivity. Consistent with these notions, employer surveys in the context of race and nationality document that employers routinely make group associations informed by their experience (Pager and Karafin, 2009; Birkelund et al., 2020).[2]

I present a model that captures these intuitive insights and highlights their implications for discrimination. In a dynamic setting, employers have noisier initial information on one group's productivity relative to another (Lundberg and Startz, 1983; Lang, 1986; Cornell and Welch, 1996; Morgan and Várdy, 2009) and trade off learning about their productivity against current-period profit maximization.[3] Part of the information observed through hiring is privately-observed by the hiring employer (Waldman, 1984; Kahn, 2013; Ge et al., 2020), such that their own hiring history influences their subsequent hiring and learning. Positive experiences with workers from a group create positive biases about their productivity, which endogenously correct themselves by leading to more hiring and learning. Negative experiences, however, create negative biases which persist by decreasing hiring of the group and therefore learning.[4] Differential learning across employers results in heterogeneous beliefs and a negatively-skewed belief distribution about the productivity of the group whose productivity is initially more uncertain.

Each period, employer beliefs about group productivity determine market clearing wages, pinned down by the marginal employer's beliefs. Optimal hiring therefore follows a cutoff rule in beliefs: employers below the cutoff do not hire from the group, preserving their negative biases. The model's key prediction is that, over time, the skewness in the belief distribution can cause the wage of the group about whose productivity employers have noisier initial

[2] For example, Pager and Karafin (2009) document the following response of an employer to a negative experience with a black worker: "You know, everyone has a couple of bad hires. And you remember those very vividly. And who that person is can really impact. That person just stuck in my head. [...] And I could see her. It was hard to not see her in other people that you meet."

[3] The complex trade-off that firms face between exploration and extraction has long been recognized as a key element of organizational learning (March, 1991) and a growing body of research combines insights from bandit problems with statistical discrimination in other contexts (Che et al., 2019; Bardhi et al., 2020; Bergman et al., 2020; Fershtman and Pavan, 2020; Komiyama and Noda, 2020). See Bergemann and Välimäki (2008) for a review of bandit problems in economics.

[4] The dynamic decision problem I study has intuitive similarities with self-confirming equilibrium models for non-cooperative games (Fudenberg and Levine, 1993a; 1993b). Both study the outcome of a learning process in which agents learn from their experiences, beliefs are not contradicted along the equilibrium path, and issues arise from insufficient learning. My model focuses on learning about the environment rather than other players' strategies, showing that some employers optimally stop learning.

information to fall and remain below their expected productivity in the long run. The model predicts discrimination due to uncertainty, even with equally productive worker groups and without prior biases or endogenous worker responses.[5] Further, since discrimination arises endogenously from expected profit maximization, it can survive competition in the form of a higher market exit rate for biased employers when new entrants also face a similar learning problem. In summary, heterogeneous biased beliefs persist within a statistical discrimination framework; they are not necessarily eliminated by learning or competition.

Like taste-based discrimination, my model generates differences between average performance and average pay of a group. In fact, it generates steady state predictions analogous to Becker (1957), with endogenous beliefs replacing preferences. Apparent taste-based discrimination can result from "incorrect" statistical discrimination, providing a new way to understand prejudice as the result of experiences shaping beliefs in distortionary ways. Biased beliefs in my model still differ starkly from a preference. They lead to distinct dynamic predictions and implications for welfare and policy, while highlighting that insights of prejudice-based models for labor market discrimination can be generated from uncertainty, without reliance on a utility function or biased updating.[6]

Endogenous group learning differs from previous work on biased beliefs and stereotypes. I study how individual biased beliefs arise when employers conduct inference on a selected sample of observations about worker group productivity, complementing work on biased beliefs creating discrimination from true group differentials (Bordalo et al., 2016), biased updating (Sarsons, 2019), worker evaluation and supervision (Bartoš et al., 2016; Glover et al., 2017), or implicit group associations (Bertrand et al., 2005). The model provides a rationale for how even employers who may be willing to give workers from any group a fair chance can develop persistent negative biases about the productivity of some groups.

A statistical discrimination framework with predictions analogous to taste-based discrimination provides a distinct way to think about discrimination. It stresses that prejudice

[5] Arrow (1973) mentions that biased priors could lead to a self-fulfilling prophecy if employers ignore subsequent information or worker responses confirm employer beliefs, but these models have no learning. I propose a distinct mechanism through which biased beliefs create discrimination without biased priors, deviation from expected profit-maximization, or endogenous worker investments.

[6] Individuals appear quick to form beliefs about groups and act on these in a way that shapes future views, consistent with the notion of prejudice from psychology (Bertrand and Duflo, 2017). My model shows 1) how biases can micro-found the reduced-form notion of prejudice in economics and 2) how biases affect decision-making in statistical discrimination models.

and statistical discrimination are not necessarily unrelated or mutually exclusive, which has a wide range of implications for studying the source of discrimination (Bohren et al., 2019a; 2019b). My model is consistent with growing empirical evidence on endogenous employer learning about groups (Leung, 2017; Lepage 2021, Benson and Lepage, 2021) and suggests that learning about some groups can be particularly slow, complementing work focusing on learning about individuals within groups (Farber and Gibbons, 1996; Lange, 2007; Arcidiacono et al. 2010; Kahn and Lange, 2014). These two implications lead to distinct policy implications from previous models. For example, my model provides a new lens to analyze policies like affirmative action, which can induce employer learning by increasing minority hiring and improve outcomes as reported in Miller (2017). In contrast with classical theories, but consistent with evidence reviewed in Lang and Kahn-Lang Spitzer (2020) as well as the large literature on contact theory reviewed in Pettigrew and Tropp (2006) and Paluck et al. (2019), my model also predicts that providing information on groups that is credible at the individual level can mitigate discrimination, as can encouraging intergroup interactions.

The paper is organized as follows: Section 1 presents the model and extensions. Section 2 situates the model in the theoretical literature. Section 3 discusses related evidence as well as implications for empirical tests and policy. Section 4 concludes.

1 Labor Market Model

1.1 Employer Information and Beliefs

Consider a large number of employers hiring workers from two observably different groups A and B (e.g. race or gender). The key feature is that, through hiring, employers learn about the productivity of worker groups, which may differ across groups for example due to historical or social factors. Assume that employers know the productivity distribution of group A, but are initially uncertain about that of group B. The important assumption is that initial information about group B's productivity is noisier, but assuming complete information on group A simplifies the analysis and exposition. Information asymmetries across worker groups are a common feature in the literature, with the distinction that I focus on the dynamic implications of an initial asymmetry for hiring and learning (Lang, 1986; Cornell and Welch, 1996; Morgan and Várdy, 2009; Lang and Manove, 2011).[7]

Each individual worker, from either group, has productivity drawn from $X \sim N(\mu, 1/\tau)$.[8] For simplicity, assume that employers know the variance $1/\tau$ and that it is equal across groups. Employers know that group A's mean productivity is μ and have common priors about the mean productivity of group B, $\mu_B \sim N(\mu_0, 1/\tau_0)$.[9] I focus on the case where $\mu_0 = \mu$, such that employers have unbiased priors. Each employer hires one worker per period, uses their hiring experiences with group B to update their beliefs, and the match dissolves after each period.[10]

In the baseline model, I make three simplifications relating to hiring and learning. First, employers observe no individual signal of productivity prior to hiring; they rely solely on group membership to predict the productivity of a worker. Second, worker signals of productivity are private and only available through an employer's own hiring. Third, there is no human capital investment or signaling by workers. Each worker is endowed with a fixed productivity and inelastically provides a unit of labor each period. The implications of each simplification are discussed in Sections 1.8-1.9.

Workers hired from group B determine the information set of employer j, S_{jt}, composed of one private signal drawn from X for each hire. The cumulative number of signals employer j has observed by time t is $K_{jt} = \sum_{n=1}^{t} \mathbb{1}(L_{Bnj} = 1)$, where L_{Bnj} is an indicator variable for whether a group B worker was hired in period n. Employers form posterior beliefs about the mean group B productivity according to the Normal updating formula

$$\mu_B | S_{jt} \sim N\left(\frac{\tau_0\mu_0 + \tau\sum_{i=1}^{K_{jt}} x_i}{\tau_0 + \tau K_{jt}}, \frac{1}{\tau_0 + \tau K_{jt}}\right). \tag{1}$$

Letting $E[\mu_B|S_{jt}] = \frac{\tau_0\mu_0+\tau\sum_{i=1}^{K_{jt}} x_i}{\tau_0+\tau K_{jt}}$ and $\text{Var}(\mu_B|S_{jt}) = \frac{1}{\tau_0+\tau K_{jt}}$, employers form posterior

[7] Information asymmetry could arise in a majority and minority setting where employers naturally observe more information about the majority group over time. It could also arise if employers, for example from group A, have better information about workers of their group due to previous experiences and interactions.

[8] Appendix 3 extends the results to more general productivity distributions.

[9] Employers have misspecified beliefs, in the sense that groups are equally productive and the true mean productivity of group B μ is a fixed constant, but employers treat it as a random variable due to uncertainty.

[10] One-period contracts focus attention on group learning by studying employers repeatedly choosing between groups. Multi-period contracts may slow down learning, but do not change relative incentives to hire and learn about group B, determined by μ_B.

beliefs about group B productivity $X_B \sim N(E[\mu_B|S_{jt}], \mathrm{Var}(\mu_B|S_{jt}) + 1/\tau)$.[11]

1.2 Hiring Decision

Consider a frictionless labor market which clears each period. I first consider a model with infinitely-lived employers learning about one cohort of workers, abstracting from product-market competition through dynamic entry and exit of firms. Employers are risk neutral, wage-takers, and maximize the present value of lifetime profits. They consider the value of learning about group B productivity, leading to a dynamic optimization problem. An individual employer's posterior beliefs are characterized by $\psi_{S_{jt}} = \{E[\mu_B|S_{jt}], \mathrm{Var}(\mu_B|S_{jt})\}$ and Ψ_t is a list of posterior beliefs across employers. Group A's wage, w_A, is time-invariant and equal to their expected productivity μ. Group B's wage, $w_{Bt}(\Psi_t)$, is set competitively across employers through market clearing each period and evolves under the influence of Ψ_t. The current-period employer payoff from hiring a worker is equal to their hire's productivity, x_i, with expected value μ for group A and $E[\mu_B|S_{jt}]$ for group B. Conditional on beliefs and wages at time t, employer j hires from group A or B to maximize their expected profits

$$
\begin{aligned}
V(\psi_{S_{jt}}, w_{Bt}(\Psi_t)) = \mathrm{Max}\{&\mu - w_A + \beta E_t[V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))], \\
&E_t[\mu_B|S_{jt}] - w_{Bt}(\Psi_t) + \beta E_t[V(\psi^{'}_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))]\}
\end{aligned}
\tag{2}
$$

where β is a discount factor. The continuation value $V(\cdot)$ includes updated beliefs $\psi^{'}_{S_{jt+1}}$ when a group B worker is hired and $\psi_{S_{jt+1}} = \psi_{S_{jt}}$ otherwise. $E_t[V(\psi^{'}_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] \geq E_t[V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))]$ since hiring from group B yields information which cannot decrease expected profits.

Endogenizing group B's wage is key because it is an outcome of interest and because intuition suggests that it should act as a counterbalancing force to bias. If the group's wage falls as a result of employers developing negatively-biased beliefs, then group B becomes cheaper, which should in turn induce employers to hire them and learn, correcting biases. I study hiring and learning decisions which account for these endogenous wage adjustments.

[11] While the true variance in productivity $1/\tau$ is known, the posterior variance of X_B is larger since employers are uncertain about the mean, increasing expected variance. Formally, the variance is given by $\int \phi_{\mu_B|S_{jt}}(m) \int \phi(x|m)(x - E[\mu_B|S_{jt}])^2 dxdm = \mathrm{Var}(\mu_B|S_{jt}) + 1/\tau$.

Optimal hiring in the current period is determined by contrasting expected profits hiring from group B versus A. The difference is positive whenever

$$\beta E_t[V(\psi'_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1})) - V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] > \mu - E_t[\mu_B|S_{jt}] - (w_A - w_{Bt}(\Psi_t)). \tag{3}$$

Equation (3) compares the expected learning value from a group B hire on the left with expected foregone profit on the right. The perceived value of learning depends on the likelihood that it will lead to changes in hiring and higher expected profits. It is maximized at $\mu_B = \mu$ since information is likeliest to affect subsequent hiring and decreases as μ_B becomes biased away from μ. In the case of negative bias, group B becomes less attractive from both a learning and a production standpoint. Thus, when prior experience suggests that group B is less productive, there is a trade-off between expected learning benefits and expected foregone profits from hiring less productive workers. This trade-off can be represented by a one-armed bandit problem, in which employers repeatedly choose between a “safe” arm (Group A) which yields a payoff from a known distribution and a “risky” arm (Group B) with an unknown payoff distribution. Obtaining comparatively low payoffs from the risky arm eventually leads the employer to stop experimenting and choose the safe arm, with the important distinction that wages and therefore payoffs are endogenous in my model.

1.2.1 Employer Learning from the Wage

One consideration is whether employers learn about the productivity of group B from the evolution of their wage. In the baseline model, I rule this out by assuming static wage expectations: employers expect the wage next period to be equal to the current one, $E[w_{Bt+1}|S_{jt}] = w_{Bt}$, keeping the model tractable since employers do not form beliefs about the beliefs of other employers. The wage in theory does carry information relevant to the learning problem faced by employers. Yet, in practice, this assumption appears particularly mild given the complexity of the problem faced by employers.

Market clearing wages summarize many private decentralized decisions that depend on factors unobserved by any given employer, rather than an aggregate price signal. Even if employers observe some relevant wage information and can invert the pricing and belief-updating processes, relative wages in practice are a function of many factors (changing skill and education, macroeconomic shocks, shifts in industry and occupation mixes, demographics, etc.), such that routinely isolating the impact of changing subjective employer beliefs about group productivity on residual wages appears implausible. Economists themselves have had long-standing unresolved debates about characterizing and decomposing wage gaps into components related to discrimination (Lang and Lehmann, 2012). Recent work on financial markets also assumes that agents neglect the informational content of prices, supported by extensive evidence on voting, trading, investing, and auctions (Eyster et al., 2019), and recent developments in modeling firm behavior surveyed in Aguirregabiria and Jeon (2019) focus on how uncertainty and learning in complex environments can lead firms to have biased beliefs, for example about demand, costs, or the behavior of other firms.

Overall, taking the current wage as a prediction for the wage next period seems like a reasonable approximation in the context of the model, especially since it will be correct in the long run. Nevertheless, I consider an extension in which employers noisily learn from outside sources such as other employers or wages in Section 1.8.

1.3 Hiring Cutoff and the Group B Wage

Define λ_{jt} as the relative willingness to pay (WTP) of employer j for a group B worker

$$\lambda_{jt} = \beta E_t[V(\psi'_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1})) - V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] - (\mu - E_t[\mu_B|S_{jt}]).$$

The trade-off between learning and foregone profit, ignoring wage considerations, is captured by λ_{jt} . It can be positive even if $E[\mu_B|S_{jt}]$ falls below μ, highlighting that employers may hire from group B even if they believe them to be less cost-effective to avoid future losses from incorrect beliefs.

Each period, labor market clearing implies that, at current wages, the fraction of employers who prefer to hire from group B is equal to the fraction of workers from the group. The group B wage each period is thus determined by the marginal employer m: the employer with the lowest λ_{jt} who must hire from the group to clear the market. Specifically, the wage is set such that the marginal employer is indifferent between hiring from either group, $\lambda_{mt} = w_{Bt}(\Psi_t) - w_A$, characterizing the optimal hiring strategy of employers stated in Proposition 1.

Proposition 1 (Optimal Hiring)
The optimal hiring strategy of employers follows a cutoff rule where employer j hires from group B at time t if and only if $\lambda_{jt} \geq \lambda_t^c$. Moreover, $\lambda_t^c = w_{Bt}(\Psi_t) - w_A$.

Proof: See Appendix 1.

Proposition 1 characterizes the cutoff below which it is optimal for employers to avoid hiring from group B at a given market wage, preserving their beliefs about the group's productivity. Since the wage gap is determined by $\lambda_t^c = \lambda_{mt}$, the optimal hiring decision of other employers immediately follows: those with λ_{jt} above the marginal employer hire from group B and others from group A, clearing the market. Market clearing thus implies

$$\nu_{\Psi_t}(\{\psi_{S_{jt}} : \lambda_{jt} \geq \lambda_t^c(w_{Bt}(\Psi_t))\}) = F_B \text{ and } \nu_{\Psi_t}(\{\psi_{S_{jt}} : \lambda_{jt} < \lambda_t^c(w_{Bt}(\Psi_t))\}) = F_A \tag{4}$$

where ν_{Ψ_t} is a measure over Ψ_t, F_g is the fraction of workers from group g, and each worker-employer pair has no incentive to deviate.

1.4 Equilibrium

An equilibrium is a stochastic process over beliefs and a mapping from beliefs to wages. Given a continuum of agents on each side of the market, this corresponds to a deterministic Markov process with corresponding transition functions characterized by Definition 1.

Definition 1 *An equilibrium is a Markov process with a distribution over beliefs Ψ_t evolving according to a transition function $T : \triangle\ (\mathbb{R} \times \mathbb{R}_+) \to \triangle\ (\mathbb{R} \times \mathbb{R}_+)$, a wage function w_{Bt}: $\triangle\ (\mathbb{R} \times \mathbb{R}_+) \to \triangle\ \mathbb{R}$ and an initial state $\Psi_0 \in \triangle\ (\mathbb{R} \times \mathbb{R}_+)$ such that every period:*

1. *Employers make expected profit maximizing hiring decisions following equation (2) and Proposition 1 for all $(\psi_{S_{jt}}, w_{Bt}(\Psi_t))$.*
2. *The labor market clears according to condition (4).*
3. *Employers update their beliefs:*

   a) Those with beliefs $\psi_{S_{jt}}$ such that $\lambda_{jt} < \lambda_t^c(w_{Bt}(\Psi_t))$ hold posterior beliefs $\psi_{S_{jt+1}} = \psi_{S_{jt}}$.

b) Those with beliefs $\psi_{S_{jt}}$ such that $\lambda_{jt} \geq \lambda_t^c(w_{Bt}(\Psi_t))$ hold posterior beliefs $\psi'_{S_{jt+1}}$ derived according to equation (1).

The first condition states that employers maximize their expected profits according to their Bellman equation and the optimal hiring rule. The second condition states that the fraction of employers with beliefs such that they want to hire from group B given current wages (λ_{jt} above the marginal) is equal to the fraction of workers from group B. The third condition states that employers below the hiring cutoff for group B do not update their beliefs, while those above the hiring cutoff update their beliefs based on the productivity of their hire according to Bayes' rule.

1.5 Biased Beliefs and Discrimination

As a result of the optimal hiring rule and equation (1), it is straightforward to characterize the asymptotic distribution of posterior beliefs described in Proposition 2.

Proposition 2 (Asymptotic Beliefs and Persistent Negative Biases)
As $t \to \infty$, beliefs of employers who remain above the hiring cutoff converge in distribution to μ. Others hold a range of beliefs such that $E[\mu_B|S_{jt}] < \mu$. The limiting fraction of employers with $E[\mu_B|S_{jt}] < \mu$ equals the fraction of group A workers.

Proof: See Appendix 1.

By standard Bayesian reasoning, posterior beliefs converge to the truth as the number of signals goes to infinity. On the other hand, employers below the cutoff (which implies $E[\mu_B|S_{jt}] < \mu$ in the long run given a strictly positive value of learning) do not hire from group B, preserving negative biases. In the long run, since unbiased employers hire from group B and biased employers hire from group A, the fraction of biased employers is equal to the fraction of group A workers.[12] Proposition 2 highlights that a subset of employers hold negatively-biased beliefs, even asymptotically.

[12] The Becker (1957) taste-based model requires that the fraction of prejudiced employers be at least as large as the fraction of group A workers to generate a wage gap. Both models thus require a majority of biased or prejudiced employers to generate a wage gap if group A is larger than group B. The fraction of employers with biased beliefs in my model is endogenously determined to be exactly equal to that of group A by market clearing, rather than being assumed. Widespread biased beliefs may be more plausible than widespread animus, and there is evidence that a large share of employers hold negative perceptions in the context of race (Lang and Lehmann, 2012).

Endogenous employer learning about worker group productivity generates a plausible distribution of beliefs for discrimination to arise. First, beliefs about group B's productivity exhibit sustained heterogeneity across employers. Second, differential learning across employers results in beliefs being negatively-skewed. Endogenous learning generates these features without relying on group differentials,[13] prejudice, or biased priors, providing a novel way to understand persistent, heterogeneous, negatively-biased beliefs.

The next consideration is whether biased beliefs generate discrimination in the form of a wage gap. Proposition 3 characterizes the evolution of group B's wage.

Proposition 3 (Wage Gap and Persistent Discrimination)
$w_{Bt}(\Psi_t)$ is strictly decreasing in t and converges to a constant $c < w_A$.

Proof: See Appendix 1.

The distribution of beliefs becomes negatively-skewed with time, because only negative bias can be stable. With hiring experience, supramarginal values of λ_{jt} become concentrated around 0 as $E[\mu_B|S_{jt}]$ becomes concentrated around μ. By definition, λ_{mt} lies below supramarginal values of λ_{jt} and thus eventually falls below 0, leading $w_{Bt}(\Psi_t)$ to fall below w_A. By market clearing, the wage cannot increase or remain constant with time. Given a continuum of employers, some employers just above the hiring cutoff are expected to have relatively negative hiring experiences with group B in any given period, such that their λ_{jt} fall below that period's cutoff. Then, the fraction of employers who want to hire from group B at the current wage is lower than the fraction of group B workers. The wage must thus decrease to induce employers to hire from the group and clear the market. Lastly, since beliefs are fixed asymptotically, there is virtually no updating, so the wage converges to a constant.

Since both groups are equally productive, the wage gap implies that group B is paid below their expected productivity. While the predicted wage gap depends on relative group productivity, the prediction that group B is paid below their expected productivity does not. The model thus predicts that persistent negatively-biased employer beliefs about group B's productivity arise and persist endogenously through selected hiring interactions and generate persistent discrimination against the group.

[13] This distinction is important even when it is unlikely that two groups have equal productivity in practice, since it predicts that closing productivity gaps would not necessarily eliminate discrimination.

1.6 Entry, Exit and Competition

A common view is that market competition should drive out biased beliefs and therefore resulting discrimination, at least in the long run. To investigate this, I augment the model with dynamic employer entry and exit from the market. The fundamental intuition regarding differential learning across employers and therefore biased beliefs remains, but exit provides a straightforward reduced-form way to introduce competition through differential exit rates based on beliefs.

Employers exit the market and are replaced with new employers at an expected aggregate rate δ each period. The exit rate influences the expected duration in the market, learning incentives, and available time for employers to potentially correct their biases. It can also directly affect the belief distribution by introducing new employers who hold different beliefs on average. I assume that employers enter with unbiased priors,[14] although Appendix 3 shows that discrimination can be amplified when priors are influenced by experienced employers.

Profit maximization is given by

$$
\begin{aligned}
V(\psi_{S_{jt}}, w_{Bt}(\Psi_t)) = \mathrm{Max}\{&\mu - w_A + (1-\delta)\beta E_t[V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))], \\
&E_t[\mu_B|S_{jt}] - w_{Bt}(\Psi_t) + (1-\delta)\beta E_t[V(\psi'_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))]\}.
\end{aligned}
$$

The exit rate of an employer should depend on profits and therefore hiring decisions determined by $E_t[\mu_B|S_{jt}]$. Since $E_t[V(\psi'_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] \geq E_t[V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))]$, employers who hire from group B earn higher expected profits of at least $w_A - w_{Bt}$ each period. Given a lower wage and equal productivity for group B, these employers are more profitable and accordingly should have a lower market exit rate, $\delta_B < \delta_A$ with $\delta = \delta_B F_B + \delta_A F_A$. If the only determinant of market exit is beliefs about the productivity of group B

[14] Prior variance may decrease if employers learn from previous cohorts of employers. This is unlikely to eliminate the initial information asymmetry since it would require employers to completely ignore their experiences, going against evidence on asymmetric employer learning (Waldman, 1984; Kahn, 2013; Ge et al., 2020), evidence surveyed in Moore et al. (2015) and Guenzel and Malmendier (2020) that decision-makers typically put too much weight on their own information and experience, and the idea that the learning problem in practice changes across worker cohorts. The relative education and experience of women and minority workers compared to that of white men was not the same in 1990 as it is today, and employment contexts have changed substantially.

($\delta_B = 0$), a differential exit rate eliminates discrimination at least in the limit.[15] Yet, firm survival in a market depends on many factors, such that firms who hire from group B also exit the market and biased beliefs may often not be pivotal (Audretsch, 1991; Schary, 1991; Black, 1995; Hellerstein et al., 2002).

In that case, the key point is that biased beliefs are not a primitive of the model, but arise endogenously through hiring experience. Therefore, as some employers held unbiased priors but developed biased beliefs through hiring, so may new employers. In the aggregate, biased beliefs and the wage gap are not necessarily eliminated by competition. A wage gap can be sustained asymptotically even if employers who hire from group A are driven out at a higher rate, as summarized in Remark 1.[16]

Remark 1 (Persistent Discrimination with Market Competition)
For some values of δ_A and δ_B with $\delta_A > \delta_B$, there exists a period $\bar{t}$ in which $w_{Bt}(\Psi_t)$ falls below w_A, remains below for all $t > \bar{t}$, and converges to a constant $c < w_A$.

Remark 1 is illustrated through simulation in the next subsection. The main difference with Proposition 3 is that the existence of a wage gap depends on parameters. At one extreme, for exit rates near zero, the existence of a wage gap directly follows from Proposition 3. At the other extreme, for very high exit rates, it is possible to introduce enough new employers with unbiased priors to hire all of group B each period, eliminating the wage gap. At the intensive margin, higher competition reduces the magnitude of the wage gap as shown in Appendix 2 and consistent with empirical evidence (Ashenfelter and Hannan, 1986; Black and Strahan, 2001). At the extensive margin, competition may not eliminate discrimination arising from endogenous biased beliefs.

1.7 Simulations

To illustrate the model's dynamics, a set of simulations was computed over 1,000 periods with 10,000 employers and 10,000 workers, 25% of which are from group B. Simulation details are

[15] Beliefs of employers above the hiring cutoff for group B converge to the group's true productivity, so an arbitrarily small mass of new entrants with $\lambda_{jt} \geq 0$ guarantees that w_{Bt} is not below w_A.

[16] In taste-based models, firm growth is important since prejudiced firms may remain in the market earning lower profits to indulge in their taste for discrimination. Then, discrimination is mitigated because unprejudiced firms grow more quickly. In my model, firms are not willing to accept a lower return for their mistaken beliefs, so growth is not conceptually necessary for discrimination to be competed away.

outlined in Appendix 2 along with additional results. Because the simulated market is finite, the evolution of beliefs and wages is stochastic rather than deterministic. Emphasis should be put on the model dynamics characterized by Propositions 1-3 and Remark 1, which do not substantively vary with parameter choice, rather than specific values of the wage gap.[17]

Panel A of Figure 1 shows the evolution of beliefs for key moments of the distribution, without entry and exit. The 25% of employers with the highest valuation for group B each period hire them and learn, so their beliefs converge towards the group's true mean productivity normalized at 0, while those of other employers are negatively biased and do not evolve. Panel B shows that the group B wage initially lies above the marginal employer's beliefs due to the value of learning, but eventually falls and remains below zero (also normalized as the group A wage) as beliefs fall below μ and the value of learning falls. With a finite market, there is a separation in the WTP of employers above and below the cutoff, seen in Panel A between the 75th and 76th percentiles. The market clearing wage can lie anywhere between these two percentiles, while the latter determines the wage with a continuum of employers as characterized in Proposition 3. If match surplus is allocated to employers, the wage is also set by the 76th percentile with a finite number of employers, as shown in Panel B.

Similarities and differences between the simulated wage path and empirical wage trends naturally do not provide a test of the model. Empirical trends depend on many sources of wage differentials outside of the model, while simulated trends depend on assumptions on priors and relative productivity, among others. For example, Appendix 2 shows that negatively-biased priors can generate a group B wage which starts and remains below that of group A, but increases over time. An analogous argument can explain the seemingly odd model prediction that employers begin by hiring group B most often and gradually decrease their hiring of the group, rather than potentially the other way around.

Figure 2 presents simulations with market entry and exit: a 2% aggregate exit rate each period and a 25% higher exit rate for employers below the hiring cutoff. The set of employers in the market is expected to be jointly replaced 3 to 4 times over the period, so the pattern is simply repeated beyond. One notable difference is that, since all employers exit the market in finite time, some employers above the hiring cutoff always have negatively-biased beliefs.

[17]Similarly, the initial state in which employers enter the market exhibits theoretically intuitive features, but is of limited practical interest. Given all employers entering simultaneously with unbiased priors, the initial group B wage may be higher than that of group A because of market clearing, but this depends on prior beliefs, relative uncertainty and productivity across groups, and potential ambiguity aversion.

There is thus a sense in which entry and exit can actually help sustain a wage gap by preventing belief convergence.

1.8 Outside Learning

In many cases, labor markets may provide few salient signals to an employer who has formed beliefs based on their own experience, and it's unclear what form outside information would need to take to be credible to an employer at the individual level. Even at similar firms, there is mismatch between employment contexts: it may be difficult for an employer to learn about the productivity of group B from observing competitors when hiring decisions and performance depend on many factors. Previous work shows that employers have better information on their hires than other employers do (Waldman, 1984; Kahn, 2013; Ge et al., 2020). A large literature surveyed in Moore et al. (2015) and Guenzel and Malmendier (2020) also documents that employers have a tendency to over-weight their own information and experience, while Benson and Lepage (2021) reports in the context of a large national retailer that a manager's hiring of black workers is influenced by their own previous hiring experiences with the group, but not those of other managers even within the same store.

Still, to the extent that employers observe some noisy information about group B's productivity outside of their own hiring, such as the hiring decisions or outcomes of a competitor, the performance of group B in other settings, or the evolution of wages, then they may learn without hiring. Such outside learning can mitigate or exacerbate bias, but has limited impact on the model's key qualitative predictions.

Consider a benchmark case in which employers get one outside signal about group B productivity per period irrespective of hiring. Outside signals are distributed $O \sim N(\mu, 1/\tau_o)$. Posterior beliefs are given by

$$\mu_B | S_{jt} \sim N\left(\frac{\tau_0\mu_0 + \tau\sum_{i=1}^{K_{jt}} x_i + \tau_o \sum_{m=1}^{t} o_m}{\tau_0 + \tau K_{jt} + \tau_o t}, \frac{1}{\tau_0 + \tau K_{jt} + \tau_o t}\right). \tag{5}$$

As long as employers put nonzero weight on their own signals, which seems like a particularly mild condition given evidence discussed above, then those who hire from group B still learn faster than those who don't if they observe both private and outside signals. The belief distribution remains negatively-skewed in any finite period, and the bias-generating mechanism at the least slows down learning. Slowing down learning itself has non-negligible implications. Statistical discrimination generally predicts that the market immediately learns equilibrium worker group productivity. One criticism is that learning is "too fast" for these models to be important in the long run (Lang and Lehmann, 2012). My model explains why learning about some groups may be particularly slow and create discrimination along the equilibrium path, reducing the lifetime income of these groups.

In the long run, if beliefs converge over time, then the wage gap is eliminated. If beliefs do not fully converge, for example because there is market entry and exit or the learning problem evolves over time, then the wage gap can remain. In practice, these two conditions are clearly satisfied. Firms, employers, and recruiters regularly enter and exit the market with finite information sets, and the relative productivity of worker groups has been evolving with changes in demographics and education, among other factors. Accordingly, an intuitive interpretation of the model is a cohort of employers learning about a cohort of workers, with imperfect transfer across cohorts. Remark 2 summarizes this result for the case of market entry and exit, which again follows from Proposition 3.

Remark 2 (Persistent Discrimination with Outside Learning)
For some values of τ_o, τ, δ_A, and δ_B with $\delta_A > \delta_B$, there exists a period $\tilde{t}$ in which $w_{Bt}(\Psi_t)$ falls below w_A, remains below for all $t > \tilde{t}$, and converges to a constant $c < w_A$.

Moreover, outside information also poses potential challenges. For instance, making hiring outcomes public within employer networks does not conceptually solve the issue that employers learn too little, because it lowers incentives for employers to hire group B and learn from their own signals, leading to free-riding (Keller et al., 2005; Hoelzemann and Klein, 2018). Equation (5) also assumes that outside signals are unbiased and unrelated to existing bias. Otherwise, outside signals could preserve or exacerbate biased beliefs (DeGroot, 1974; Gentzkow and Shapiro, 2006; Baliga et al., 2013; Enke and Zimmermann, 2017; Fryer et al., 2018). In any case, outside learning suggests two additional implications. First, discrimination may differ across settings based on the observability of competitors, workers, wages, and output. Second, there is potential scope for the provision of information.

1.9 Other Modeling Features

1.9.1 Firm Size and Hiring Policy

While the model assumes that each employer only hires one worker per period, larger employers who hire more workers have a higher value of learning and may learn more quickly. Negative biases may be less likely to arise and persist, and these employers may hire a higher fraction of group B workers, consistent with evidence reported in Holzer (1998) and Miller (2017) for black workers. These implications presumably relate to large establishments with centralized, professional human resources (HR) departments rather than large firms with decentralized hiring across smaller establishments, with the latter having been associated with increased discriminatory outcomes (Berson et al., 2019). When parts of the hiring process is decentralized, individual managers play an important role in the racial composition of hires (Giuliano et al., 2009; Giuliano and Ransom, 2013; Benson, Board and Meyer-ter-Vehn, 2019; Benson and Lepage, 2021), information aggregation across different local markets is nontrivial, and common HR policies like pre-employment testing and hiring algorithms typically fail to address concerns of endogenous employer learning specifically (Hoffman, 2018; Bergman et al., 2020; Benson and Lepage, 2021).

Implications for market-level discrimination in the model remain limited if each establishment hires a negligible fraction of the labor force and there is size heterogeneity above the hiring cutoff. Unless all of group B is hired by large establishments with centralized hiring, then these establishments are not marginal, by definition, and the wage is determined by smaller establishments who learn more slowly. Casual empiricism certainly suggests that some small firms and large firms with decentralized hiring hire workers from groups typically of interest in the discrimination literature. A back of the envelope calculation suggests that around 17% of black workers were employed at firms with less than 25 workers in 1998, and this proportion is substantially larger for establishments under 25 workers (Headd, 2000).

1.9.2 Signals of Individual Productivity and Endogenous Worker Investments

Consider the case in which employers observe a noisy signal s_i of individual worker productivity x_i at the hiring stage and do not rely solely on group membership g to predict productivity. This signal is exogenous, rather than the result of an investment choice, and can be thought of as a score on a pre-employment test. Employers observe

$$s_i = x_i + \varepsilon_i$$

where $\varepsilon_i \sim N(0, 1/\tau_\varepsilon)$ is i.i.d. random noise. They estimate productivity according to the following rule

$$E[x_i|s_i, S_{jt}] = \gamma s_i + (1-\gamma)E[\mu_g|S_{jt}]$$

where $\gamma_{gjt} = \frac{1/\tau + Var[\mu_g|S_{jt}]}{1/\tau + Var[\mu_g|S_{jt}] + 1/\tau_\varepsilon}$ is a measure of the signal's precision. Negatively-biased beliefs about the mean productivity of group B arise as in the baseline model. Since employers above the hiring cutoff are willing to pay more for a group B worker conditional on a given signal value, workers and employers sort such that hiring and learning dynamics are also unchanged. Workers can be indexed by their signal value, with the same learning problem arising for each worker "type" and a market-clearing wage for each type-group pair.

Discrimination may still vary by occupation, skill, and education depending on the variance in productivity and productivity signals. These variances determine the extent to which employers rely on group membership to predict productivity, and therefore the importance of the learning problem. Discrimination empirically appears smaller for high-skill workers, at least in the case of race (Lang and Lehmann, 2012). Differences in the information available at the time of hiring, variance in productivity, or the speed with which the market learns individual worker productivity, could all help explain this empirical regularity (Arcidiacono et al., 2010; Lindqvist and Vestman, 2011).

When groups are ex-ante equally productive, statistical discrimination models usually generate outcome disparities by showing that workers from group B may face different incentives to invest in human capital, for example due to employers perceiving their signals of productivity as noisier (Lundberg and Startz, 1983) or because they hold negative stereotypes against them (Coate and Loury, 1993). Statistical discrimination therefore arises when group B becomes less productive due to lower investment.

While a formal model of endogenous worker investment is beyond the scope of this paper, in my model, even if employers have biased beliefs on average, workers and employers sort such that group B is hired by employers above the cutoff who have approximately unbiased average beliefs with experience. Accordingly, group B doesn't necessarily have incentives to invest differentially in human capital due to biased beliefs of employers. Nevertheless, group

B may expect a different return for the same investment if relative wages across investment levels vary due to the nature of individual signals of productivity. Group B workers may be incentivized to sort into areas or occupations where the information asymmetry problem faced by employers is lesser, providing a rationale for group specialization. Similarly, if group B earns lower returns from the labor market overall, they may have incentives to invest less in human capital which could exacerbate discrimination.

2 Relationship with Other Theories

The model generates steady state predictions analogous to those from Becker (1957), with preferences replaced by endogenous beliefs:

- An employer hires group A if the wage gap is smaller than λ_{jt} and group B otherwise.
- If enough employers have (approximately) correct beliefs to hire all of group B, there is effective segregation without a wage gap.
- If enough employers have biased beliefs, there is a wage gap determined by the marginal employer.

The model has intuitive similarities with taste-based discrimination, namely a difference between average productivity and average pay of a group, but without deviating from a statistical discrimination framework. This is a key point given that taste-based discrimination has often been criticized for the arbitrariness of including preferences in a utility function. The important insights of prejudice-based models for labor market discrimination do not in fact rely on preferences, but can be understood as arising from uncertainty. Biased beliefs capture context-dependent aspects such as gender-based discrimination and differentials by skill and education, which are less compatible with the notion of an aversion to contact. Widespread biased beliefs may also be more plausible than widespread overt animus, which evidence suggests has been steadily decreasing over past decades, unlike outcome differentials (Lang and Lehmann, 2012). This does not imply that preferences and biased beliefs are necessarily substitutes, because they differ fundamentally in how discrimination arises, evolves, and can be mitigated.[18] Indeed, studying how individual discriminatory responses evolve over time and with experience is a key avenue to test belief-based discrimination and distinguish it from other sources, as exemplified in the next section.

The model complements the statistical discrimination literature by relaxing the assumption that employers have correct equilibrium beliefs about group productivity and instead modeling learning. In many contexts, the assumption that employers know the productivity of worker groups or instantly learn it in equilibrium seems implausible, yet little work considers how relaxing the assumption can have important implications.[19] Discrimination caused by biased beliefs can arise without grounds for classical statistical discrimination. It does not arise from objective market-level group differences, but subjective, potentially flawed beliefs. It is not a self-fulfilling prophecy nor the result of coordination failures between firms and workers. The discriminated-against group cannot be seen as having "played a hand" in justifying discrimination against them, and discrimination can be sustained without prior bias or homogeneous beliefs.[20]

Another point concerns efficiency and equality. In statistical discrimination models, outcomes usually reflect true average productivity, so ending discrimination may not help group B on average. As a result, this type of discrimination is generally regarded as efficient. In my model, workers are paid below their expected productivity because of what are essentially employer mistakes. A social planner concerned with inequality or equality of opportunity could improve group B outcomes at no efficiency cost through increased employer learning.

[18] If biased beliefs are reinforced through behavioral primitives in the utility function, they could be essentially indistinguishable from a taste. Individuals with a taste for discrimination may gather and interpret information in a way that validates and justifies their prejudice (Nickerson, 1998).

[19] Aigner and Cain (1977) state in their model that group means "are estimated without bias" by employers and that "as an explanation of discrimination against blacks, a theory of discrimination based on employers' mistakes is even harder to accept than the explanation based on employers' 'tastes for discrimination,' because the 'tastes' are at least presumed to provide a source of 'psychic gain' (utility) to the discriminator."

[20] The homogeneous prior assumption usually made in the literature can be important to generate long-run discrimination. Otherwise, some employers may be better at interpreting signals (Aigner and Cain, 1977) or have more accurate priors (Coate and Loury, 1993). Other employers would learn or exit the market, such that the need for the discriminated-against group to adjust is unclear.

3 Related Evidence and Implications

Discrimination from biased beliefs is consistent with a growing body of evidence, including on endogenous employer learning about groups specifically. Leung (2017) documents on an online job board that previous hiring experiences of employers with workers from particular countries affect the subsequent likelihood of hiring workers from those countries. Lepage (2021) studies endogenous bias formation through endogenous learning using an experimental labor market. When tasked with hiring the most productive workers across several periods and faced with initial information asymmetry about the productivity of two groups, employers reduce their hiring of the more uncertain group following negative experiences, captured by hiring relatively low productivity workers, leading to reduced learning and persistent negative biases observed through belief elicitation. In contrast, positive experiences increase hiring and learning, creating a negatively-skewed distribution of beliefs across employers as predicted by the model. Benson and Lepage (2021) uses longitudinal employment records from a large US retailer to document that the hiring history of managers creates heterogeneity in their hiring of worker groups. Managers increase (decrease) their relative hiring of black and white workers following positive (negative) experiences with these groups, with proportionally larger impacts for black workers consistent with stronger updating by managers. Further, early negative experiences with black workers persistently decrease relative hiring of the group over subsequent hiring cycles, unlike early positive experiences or early negative ones with white workers. These findings are particularly consistent with hiring experiences of employers leading to belief updating about the performance of worker groups. Belief updating in turn affects subsequent hiring patterns in a manner consistent with endogenous employer learning and which systematically decreases relative hiring of minority workers. Across three different contexts, these papers support the notion that the mechanism presented in this paper substantively affects discriminatory behavior in practice.

In addition, the model generates distinct policy implications which are also consistent with a growing body of evidence. Central to the model is the idea that employers learn about groups through interaction and exposure. This feature suggests that information on group productivity can help distinguish between the different theories of discrimination: preferences should not respond to information about productivity and information on groups should not affect average outcomes if they reflect true group productivity as in classical statistical discrimination models. Yet, recent evidence surveyed in Lang and Kahn-Lang Spitzer (2020) indicates that increased information on groups mitigates discriminatory behavior, consistent with biased beliefs.

More specifically, policies which induce employers to learn more through their own experiences may be particularly effective in mitigating biased beliefs. The model provides a new lens to study policies like internships, worker subsidies, and affirmative action which induce employers to hire more workers from group B and learn, consistent with persistent improved minority outcomes even after the end of affirmative action policies, as documented in Miller (2017). Critics of affirmative action often state that the worker best qualified for a position should be hired, independent of group membership. This argument hinges on the assumption that employers know ex-ante which worker is most qualified and therefore have correct beliefs about group productivity. My model suggests that this may not be the case and that affirmative action may in fact be necessary to move towards the point where the worker best qualified for a position is hired, independent of group membership. Relatedly, Pettigrew and Tropp (2006) and Paluck et al. (2019) conclude from their surveys that intergroup contact, particularly intense collaborative exposure and integration, typically reduces prejudice. These predictions follow directly from my framework of belief updating. One historical example is World War II, often discussed as a shock through which employers learned about the productivity of women and minority groups (Goldin, 1991).

This type of discrimination also leaves scope for interventions by firms and employers. Firms in which individuals hold discretionary power in hiring decisions have incentives to eliminate biased beliefs as studied in this paper, but standard policies typically focus on reducing individual uncertainty rather than incorporating the notion of learning about worker groups.[21] One exception is Bergman et al. (2020), which studies resume screening algorithms in a setting where firms balance selecting workers from previously successful groups with selecting from under-represented groups. They find that algorithms which value learning can improve both hiring performance and diversity.

Lastly, the model and its related evidence have implications for identifying the source of discrimination in economic markets, which remains a key objective of the literature. Tra-

[21] Similarly, the proportion of minority or disadvantaged groups in management positions is low compared to that of workers (Giuliano et al., 2009; Blau and Kahn, 2017; Benson and Lepage, 2021). Evidence on how biased beliefs arise and evolve for these employers could shed light on whether increasing their representation could mitigate discrimination, potentially encouraging segregation and mitigating wage gaps.

ditionally, this has meant distinguishing between taste-based and statistical discrimination. Empirical tests often provide indirect evidence by comparing observed outcomes to those expected from true group differences, with the residual classified as taste. Such logic is conceptually inadequate, because the absence of observable productivity differentials does not imply a taste for discrimination given the potential for statistical discrimination with incorrect beliefs. Similarly, employers responding to information is consistent with statistical discrimination, but does not imply that employers hold correct beliefs on average or use information correctly. Bohren et al. (2019a) studies the empirical identification challenge posed by biased beliefs and stresses that they are rarely considered in the literature. My model provides a specific mechanism through which biased beliefs can generate discrimination which blurs the line between the two classical theories, highlighting that biased beliefs should not be ignored as a potential source of discrimination.

4 Conclusion

This paper presents a new model of discrimination in which persistent, heterogeneous employer biased beliefs about the productivity of worker groups arise and can create disparate outcomes. Given initial uncertainty about the relative productivity of worker groups, employers systematically develop biased beliefs through endogenous learning influenced by their previous hiring experiences with groups. These biased beliefs can create discrimination against worker groups whose productivity is initially more uncertain to employers, even with expected profit-maximizing employers in a competitive market with equally-productive worker groups, no prior bias or prejudice, and without endogenous worker investments.

The model generates steady state predictions analogous to Becker (1957), replacing preferences with endogenous biased beliefs and highlighting that some of what is usually classified as a taste may be understood as biased beliefs. It provides a new way to understand prejudice in the labor market as the result of selected interactions between groups distorting beliefs and behavior. It generates these novel implications while being set within a statistical discrimination framework in which learning about groups is modeled explicitly, complementing previous models in that literature. Biased beliefs in this paper arise from information frictions, with implications for understanding the relationship between theories of discrimination, empirically studying the source of discrimination, and policy.

Overall, evidence from employer surveys and the empirical literature supports the intuitive notion that individual employers make group associations which evolve over time with information and experience, influencing discriminatory behavior. This feature of discrimination is absent from classical models of discrimination, creating a gap between the theoretical literature and discrimination documented in practice. My model fills this gap by providing the first formalization of employer learning about worker groups and working out its key implications for labor market discrimination.

The model focuses on profit-maximizing employers who are Bayesian over their own experiences, although existing work documents behavioral elements which could amplify discrimination based on biased beliefs and increase the connection with preferences. This interaction is a natural direction for future research and suggests that biased beliefs in this model may constitute a lower bound on discriminatory behavior in many empirical settings.

References

[1] Aguirregabiria, V. and Jeon, J., 2019. Firms' beliefs and learning: Models, identification, and empirical evidence. Review of Industrial Organization, 56, pp.203-235.

[2] Aigner, D.J. and Cain, G.G., 1977. Statistical theories of discrimination in labor markets. ILR Review, 30(2), pp.175-187.

[3] Altonji, J.G. and Pierret, C.R., 2001. Employer learning and statistical discrimination. The Quarterly Journal of Economics, 116(1), pp.313-350.

[4] Arcidiacono, P., Bayer, P. and Hizmo, A., 2010. Beyond signaling and human capital: Education and the revelation of ability. American Economic Journal: Applied Economics, 2(4), pp.76-104.

[5] Arnold, D., Dobbie, W. and Yang, C.S., 2018. Racial bias in bail decisions. The Quarterly Journal of Economics, 133(4), pp.1885-1932.

[6] Arrow, K., 1973. The theory of discrimination. Discrimination in Labor Markets, 3(10), pp.3-33.

[7] Ashenfelter, O. and Hannan, T., 1986. Sex discrimination and product market competition: The case of the banking industry. The Quarterly Journal of Economics, 101(1), pp.149-173.

[8] Audretsch, D.B., 1991. New-firm survival and the technological regime. The Review of Economics and Statistics, 73(3), pp.441-450.

[9] Baliga, S., Hanany, E. and Klibanoff, P., 2013. Polarization and ambiguity. American Economic Review, 103(7), pp.3071-83.

[10] Bardhi, A., Guo, Y. and Strulovici, B., 2020. Early-Career Discrimination: Spiraling or Self-Correcting?. Working Paper. Northwestern University.

[11] Bartoš, V., Bauer, M., Chytilová, J. and Matějka, F., 2016. Attention discrimination: Theory and field experiments with monitoring information acquisition. American Economic Review, 106(6), pp.1437-75.

[12] Becker, G.S., 1957. The economics of discrimination. University of Chicago press.

[13] Benson, A., Board, S. and Meyer-ter-Vehn, M., 2019. Discrimination in hiring: Evidence from retail sales. Unpublished. University of Minnesota.

[14] Benson, A. and Lepage, L., 2021. The Emergence of Hiring Discrimination. Unpublished. University of Minnesota.

[15] Bergemann, D. and Välimäki, J., 2008. Bandit problems. The New Palgrave Dictionary of Economics: Volume 18, pp.336-340.

[16] Bergman, P., Li, D. and Raymond, L., 2020. Hiring as Exploration. SSRN working paper 3630630.

[17] Berson, C., Laouénan, M. and Valat, E., 2019. Outsourcing recruitment as a solution to prevent discrimination: A Correspondence Study. IZA Discussion Paper No. 12132

[18] Bertrand, M., Chugh, D. and Mullainathan, S., 2005. Implicit discrimination. American Economic Review, 95(2), pp.94-98.

[19] Bertrand, M. and Duflo, E., 2017. Field experiments on discrimination. In Handbook of Economic Field Experiments (Vol. 1, pp. 309-393). North-Holland.

[20] Birkelund, G.E., Johannessen, L.E., Rasmussen, E.B. and Rogstad, J., 2020. Experience, stereotypes and discrimination. Employers' reflections on their hiring behavior. European Societies, 22(4), pp.503-524.

[21] Black, D.A., 1995. Discrimination in an equilibrium search model. Journal of Labor Economics, 13(2), pp.309-334.

[22] Black, S.E. and Strahan, P.E., 2001. The division of spoils: rent-sharing and discrimination in a regulated industry. American Economic Review, 91(4), pp.814-831.

[23] Blau, F.D. and Kahn, L.M., 2017. The gender wage gap: Extent, trends, and explanations. Journal of Economic Literature, 55(3), pp.789-865.

[24] Bohren, J.A., Haggag, K., Imas, A. and Pope, D.G., 2019a. Inaccurate statistical discrimination (No. w25935). National Bureau of Economic Research.

[25] Bohren, J.A., Imas, A. and Rosenberg, M., 2019b. The dynamics of discrimination: Theory and evidence. American Economic Review, 109(10), pp.3395-3436.

[26] Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2016. Stereotypes. The Quarterly Journal of Economics, 131(4), pp.1753-1794.

[27] Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2019. Beliefs about gender. American Economic Review, 109(3), pp.739-73.

[28] Che, Y.K., Kim, K. and Zhong, W., 2020. Statistical discrimination in ratings-guided markets. arXiv preprint arXiv:2004.11531.

[29] Coate, S. and Loury, G.C., 1993. Will affirmative-action policies eliminate negative stereotypes?. American Economic Review, 83(5), pp.1220-1240.

[30] Cornell, B. and Welch, I., 1996. Culture, information, and screening discrimination. Journal of Political Economy, 104(3), pp.542-571.

[31] DeGroot, M.H., 1974. Reaching a consensus. Journal of the American Statistical Association, 69(345), pp.118-121.

[32] Enke, B. and Zimmermann, F., 2017. Correlation neglect in belief formation. The Review of Economic Studies, 86(1), pp.313-332.

[33] Eyster, E., Rabin, M. and Vayanos, D., 2019. Financial markets where traders neglect the informational content of prices. The Journal of Finance, 74(1), pp.371-399.

[34] Farber, H.S. and Gibbons, R., 1996. Learning and wage dynamics. The Quarterly Journal of Economics, 111(4), pp.1007-1047.

[35] Fershtman, C. and Gneezy, U., 2001. Discrimination in a segmented society: An experimental approach. The Quarterly Journal of Economics, 116(1), pp.351-377.

[36] Fershtman, D. and Pavan, A., 2020. Soft affirmative action and minority recruitment. arXiv preprint arXiv:2004.14953.

[37] Fryer, R.G., Harms, P. and Jackson, M.O., 2018. Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. Journal of the European Economic Association, 17(5), pp.1470-1501.

[38] Fudenberg, D. and Levine, D.K., 1993a. Self-confirming equilibrium. Econometrica, 61(3), pp.523-545.

[39] Fudenberg, D. and Levine, D.K., 1993b. Steady state learning and Nash equilibrium. Econometrica, 61(3), pp.547-573.

[40] Ge, S., Moro, A. and Zhu, B., 2018. Testing for Asymmetric Employer Learning and Statistical Discrimination. SSRN working paper 3243966.

[41] Gentzkow, M. and Shapiro, J.M., 2006. Media bias and reputation. Journal of Political Economy, 114(2), pp.280-316.

[42] Giuliano, L., Levine, D.I. and Leonard, J., 2009. Manager race and the race of new hires. Journal of Labor Economics, 27(4), pp.589-631.

[43] Giuliano, L. and Ransom, M.R., 2013. Manager ethnicity and employment segregation. ILR Review, 66(2), pp.346-379.

[44] Glover, D., Pallais, A. and Pariente, W., 2017. Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. The Quarterly Journal of Economics, 132(3), pp.1219-1260.

[45] Goldin, C.D., 1991. The role of World War II in the rise of women's employment. American Economic Review, 81(4), pp.741-756.

[46] Guenzel, M. and Malmendier, U., 2020. Behavioral corporate finance: The life cycle of a CEO career (No. w27635). National Bureau of Economic Research.

[47] Headd, B., 2000. The characteristics of small-business employees. Monthly Lab. Rev., 123, p.13.

[48] Hellerstein, J.K., Neumark, D. and Troske, K.R., 2002. Market Forces and Sex Discrimination. Journal of Human Resources, 37(2), pp.353-380.

[49] Hoelzemann, J. and Klein, N., 2018. Bandits in the Lab. Cahiers de recherche 2018-09, Université de Montréal, Département de sciences économiques.

[50] Hoffman, M., Kahn, L.B. and Li, D., 2018. Discretion in hiring. Quarterly Journal of Economics, 133(2), pp.765-800.

[51] Holzer, H.J., 1998. Why do small establishments hire fewer blacks than large ones?. Journal of Human Resources, 33(4), pp.896-914.

[52] Kahn, L.B., 2013. Asymmetric information between employers. American Economic Journal: Applied Economics, 5(4), pp.165-205.

[53] Kahn, L.B. and Lange, F., 2014. Employer learning, productivity, and the earnings distribution: Evidence from performance measures. The Review of Economic Studies, 81(4), pp.1575-1613.

[54] Keller, G., Rady, S. and Cripps, M., 2005. Strategic experimentation with exponential bandits. Econometrica, 73(1), pp.39-68.

[55] Komiyama, J. and Noda, S., 2020. On statistical discrimination as a failure of social learning: A multi-armed bandit approach. arXiv preprint arXiv:2010.01079.

[56] Landsman, R., 2018. Gender differences in executive departure. Unpublished. Bucknell University.

[57] Lang, K., 1986. A language theory of discrimination. The Quarterly Journal of Economics, 101(2), pp.363-382.

[58] Lang, K., Kahn-Lang Spitzer, A., 2020. Racial discrimination: An economic perspective. Forthcoming in Journal of Economic Perspective.

[59] Lang, K. and Lehmann, J.Y.K., 2012. Racial discrimination in the labor market: Theory and empirics. Journal of Economic Literature, 50(4), pp.959-1006.

[60] Lang, K. and Manove, M., 2011. Education and labor market discrimination. American Economic Review, 101(4), pp.1467-96.

[61] Lange, F., 2007. The speed of employer learning. Journal of Labor Economics, 25(1), pp.1-35.

[62] Laouénan, M. and Rathelot, R., 2017. Ethnic discrimination on an online marketplace of vacation rental. HAL-01514713 Working Paper.

[63] Lepage, L., 2021. Endogenous Belief Formation in Hiring and Discrimination. Unpublished. University of Michigan.

[64] Lesner, R.V., 2018. Testing for statistical discrimination based on gender. Labour, 32(2), pp.141-181.

[65] Leung, M.D., 2017. Learning to hire? Hiring as a dynamic experiential learning process in an online market for contract labor. Management Science, 64(12), pp.5651-5668.

[66] Lindqvist, E. and Vestman, R., 2011. The labor market returns to cognitive and non-cognitive ability: Evidence from the Swedish enlistment. American Economic Journal: Applied Economics, 3(1), pp.101-28.

[67] Lundberg, S.J. and Startz, R., 1983. Private discrimination and social intervention in competitive labor market. American Economic Review, 73(3), pp.340-347.

[68] March, J.G., 1991. Exploration and exploitation in organizational learning. Organization science, 2(1), pp.71-87.

[69] Miller, C., 2017. The persistent effect of temporary affirmative action. American Economic Journal: Applied Economics, 9(3), pp.152-90.

[70] Mobius, M., Rosenblat, T. and Wang, Q., 2016. Ethnic discrimination: Evidence from China. European Economic Review, 90, pp.165-177.

[71] Moore, D.A., Tenney, E.R. and Haran, U., 2015. Overprecision in judgment. The Wiley Blackwell Handbook of Judgment and Decision Making, pp.182-209

[72] Morgan, J. and Várdy, F., 2009. Diversity in the Workplace. American Economic Review, 99(1), pp.472-85.

[73] Moro, A. and Norman, P., 2004. A general equilibrium model of statistical discrimination. Journal of Economic Theory, 114(1), pp.1-30.

[74] Nickerson, R.S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), pp.175-220.

[75] Pager, D. and Karafin, D., 2009. Bayesian bigot? Statistical discrimination, stereotypes, and employer decision making. The Annals of the American Academy of Political and Social Science, 621(1), pp.70-93.

[76] Paluck, E.L., Green, S.A. and Green, D.P., 2019. The contact hypothesis re-evaluated. Behavioural Public Policy, 3(2), pp.129-158.

[77] Pettigrew, T.F. and Tropp, L.R., 2006. A meta-analytic test of intergroup contact theory. Journal of personality and social psychology, 90(5), p.751.

[78] Phelps, E.S., 1972. The statistical theory of racism and sexism. The American Economic Review, 62(4), pp.659-661.

[79] Reuben, E., Sapienza, P. and Zingales, L., 2014. How stereotypes impair women's careers in science. Proceedings of the National Academy of Sciences, 111(12), pp.4403-4408.

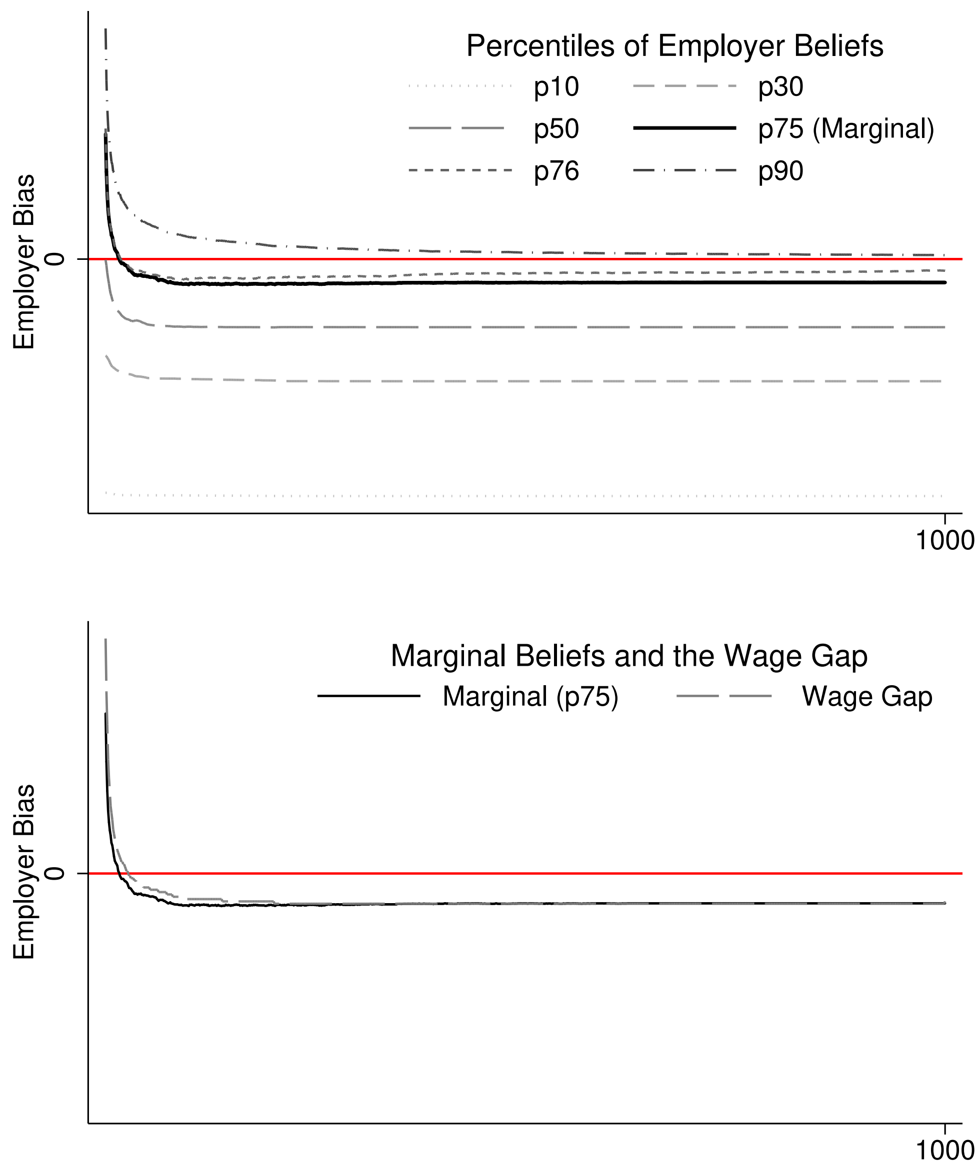
[80] Sarsons, H., 2019. Interpreting signals in the labor market: Evidence from medical referrals. Unpublished. University of Toronto.

[81] Schary, M.A., 1991. The probability of exit. The RAND Journal of Economics, 22(3), pp.339-353.

[82] Van Dalen, H.P. and Henkens, K., 2019. Do stereotypes about older workers change? A panel study on changing attitudes of managers. Forthcoming at International Journal of Manpower.

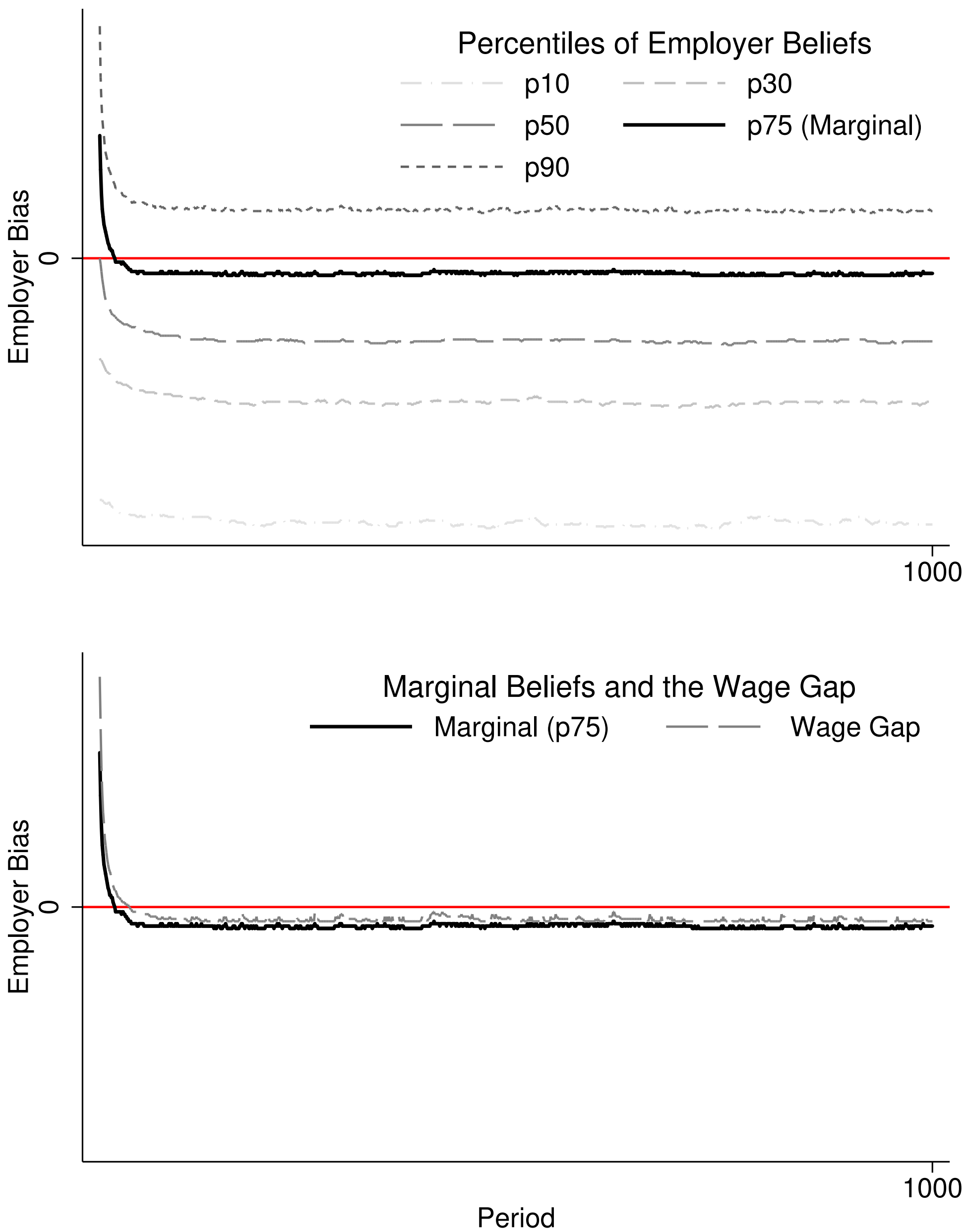
[83] Waldman, M., 1984. Job assignments, signalling, and efficiency. The RAND Journal of Economics, 15(2), pp.255-267.

Figure 1: Model Simulation without Entry and Exit



The fraction of group B workers is 0.25. Worker productivity is distributed $N(0,2)$, prior beliefs are distributed $N(0,1)$. w_A is normalized to 0 and β is set to 0.9.

Figure 2: Model Simulation with Market Entry and Exit, 25% Exit Differential



The aggregate exit rate corresponds to 2% each period, with a 25% higher exit rate for employers below the hiring cutoff for group B. New entrants have mean beliefs equal to 0 (unbiased). See Figure 1 for other parameter choices.

Appendix 1 - Proofs of Propositions 1-3

Proposition 1

By market clearing, the marginal employer is indifferent between hiring from either group, implying $\lambda_{mt} = w_{Bt}(\Psi_t) - w_A$. Define $\lambda_{mt} = \lambda_t^c$. Given current beliefs and wages, profit maximization implies that employers with $\lambda_{jt} > \lambda_t^c$ strictly prefer to hire from group B while those with $\lambda_{jt} < \lambda_t^c$ strictly prefer to hire from group A. Thus, λ_t^c represents the cutoff relative WTP for a group B worker in period t.

Proposition 2

Given the prior $\mu_B \sim N(\mu_0, 1/\tau_0)$ and i.i.d hiring signals $x_1, ..., x_K$ drawn from $X \sim N(\mu, 1/\tau)$, the Bayesian Central Limit Theorem implies under standard regulatory conditions that the posterior belief distribution converges in distribution to μ as $K \to \infty$. The posterior distribution of beliefs for employers who remain above the hiring cutoff in the long run converges in distribution to μ. For almost all of these these employers, this implies that the value of learning converges to 0 such that $\lambda_{jt} \to 0$ as $K \to \infty$. Market clearing requires that a subset of employers hire from group B, for almost all of whom $\lambda_{jt} \to 0$ as $K \to \infty$ and $\lambda_{jt} \geq w_{Bt}(\Psi_t) - w_A$. Thus, $w_A \geq w_{Bt}(\Psi_t)$ asymptotically.

Market clearing also requires that a subset of employers hire from group A asymptotically, implying $\lambda_{jt} \leq \lambda_t^C$ for those employers. Define

$$\begin{aligned}
&\Delta V_{jt} = E_t[V(\psi'_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] - E_t[V(\psi_{S_{jt+1}}, w_{Bt+1}(\Psi_{t+1}))] \text{ and} \\
&\Delta f_{jt} = \mu - E[\mu_B | S_{jt}].
\end{aligned}$$

Employer j hires from group A only if $\beta \Delta V_{jt} - \Delta f_{jt} \leq w_{Bt}(\Psi_t) - w_A$. Since the value of information ΔV_{jt} is weakly positive, then $\Delta f_{jt} > 0$ for this group. $\Delta f_{jt} > 0$ implies that $E[\mu_B | S_{jt}] < \mu$. Employers who hire from group A asymptotically must have negatively-biased beliefs. Let F_B denote the fraction of group B workers. Asymptotically, since unbiased employers hire from B and biased employers from A, the fraction of biased employers is equal to $1 - F_B$ by market clearing.

Proposition 3

First, I show that w_{Bt} is strictly decreasing in t. Define E_{Bt} as the subset of employers who hire from group B in a given period t, with the fraction of employers in E_{Bt} equaling F_B. By definition, $\lambda_{jt} \geq w_{Bt} - w_A$ for these employers. Given a continuum of employers, some employers arbitrarily close to the cutoff observe a low signal, such that there exists $e_{Bt+1} \subset E_{Bt}$ with $\lambda_{jt+1} < w_{Bt} - w_A \leq \lambda_{jt}$.[22] Suppose $w_{Bt+1} \geq w_B$, then $E_{Bt+1} \subset E_{Bt}$ and the labor market doesn't clear. Thus, w_{Bt+1} must be smaller than w_{Bt} for all t.

Second, I show that $w_{Bt} \to c \in \mathbb{R}$ as $t \to \infty$. Since w_{Bt} is strictly decreasing in t, this is equivalent to establishing that w_{Bt} cannot fall below an arbitrarily low limit $\underline{\mathrm{w}}$. In any period, even asymptotically, employers below the hiring cutoff have observed a finite number of signals (if any). Then, they have a strictly positive value of learning about group B and posterior mean beliefs strictly above negative infinity. Denote $\lambda_{\underline{j}} = \underline{\mathrm{w}} > -\infty$ where $\lambda_{\underline{j}}$ is the supremum relative WTP for a group B worker for employers below the cutoff as $t \to \infty$. Then, $w_{Bt} \geq \underline{\mathrm{w}}$ for any t. Since w_{Bt} is strictly decreasing in t but bounded below, it must converge to a constant as $t \to \infty$.

Third, I show that $c < w_A$. For any $\varepsilon > 0$, there exists a t large enough such that fraction $F_B - \varepsilon$ of employers currently hiring from Group B have value of learning smaller than ε and will hire from Group B in the limit.[23] There also exists $t' > t$ arbitrarily large such that beliefs of employers hiring from Group B at t' are almost entirely driven by signals observed between t and t'. More precisely, $\mu_B|S_{t'j}$ follows approximately the same distribution as $\mu_B|\{S_{t'j} \setminus S_{jt}\}$ with the same parameters. Given that $E[\mu_B|\{S_{t'j} \setminus S_{jt}\}]$ converges to μ for almost all employers who hire from group B, some employers who hire from group B at t' have posterior mean beliefs below μ[24] and a value of learning smaller than ε, such that their relative WTP for a group B worker λ_{jt} is below 0. By market clearing, the relative WTP of the marginal employer is no greater than the infimum relative WTP of employers hiring from group B, implying that $\lambda_{mt} = w_{Bt} - w_A < 0$ and thus that $w_{Bt} < w_A$ for $t > t'$. Since w_{Bt} is strictly decreasing in t, then $c < w_A$.

[22] This does not rely on unbounded signals. The continuum assumption ensures that a mass of employers is arbitrarily close to the cutoff each period.

[23] This is because the value of learning and the probability that an employer currently hiring from group B falls below the cutoff next period go to 0 asymptotically.

[24] The probability that the posterior beliefs of employers all converge in distribution to μ from above is 0 given a large number of employers and signals.

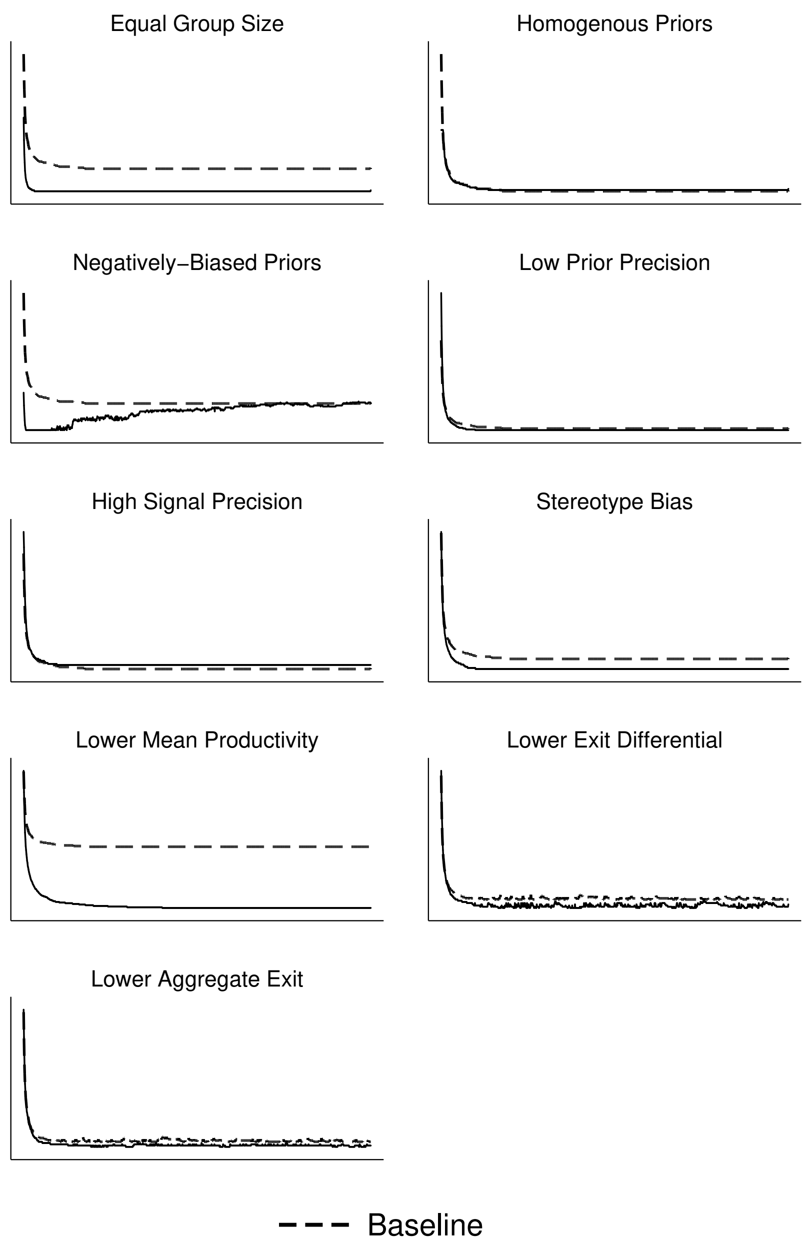
Appendix 2 - Simulations and Comparative Dynamics

Given a prior distribution of beliefs, the initial market-clearing wage when employers maximize their expected profits is found. Beliefs are updated such that those above the cutoff receive a signal of productivity from group B and others retain their beliefs. Given this new distribution of beliefs, a new market-clearing wage is found, and the process is repeated. The dynamic optimization problem is solved for a discretized state space which gives the value of learning for combinations of beliefs and wages through interpolation. Worker productivity is distributed $N(0, 2)$ and prior beliefs are distributed $N(0, 1)$. The group A wage w_A is normalized to 0 and the discount factor β is set to 0.9.

The expected size of the wage gap is influenced by the exogenous parameters of the model as in Figure A2-1. Namely, as in the Becker model, a higher fraction of group B workers is predicted to lead to a lower wage for group B. A lower mean group productivity also leads to a lower wage. If employers have negatively-biased priors about group B productivity, then their wage will be lower initially and reach a similar level in the long run. A higher prior precision or lower variance in productivity increases the wage of group B. Assuming homogeneous rather than unbiased priors has little impact on the wage (slightly higher), while introducing stereotype bias through employers overestimating the precision of their signals (or equivalently underestimating the variance in group B's productivity) decreases the wage. With entry and exit of employers, when new employers hold unbiased priors, a lower exit rate differential for employers who hire from group A leads to a lower wage for group B, as does a lower aggregate exit rate.

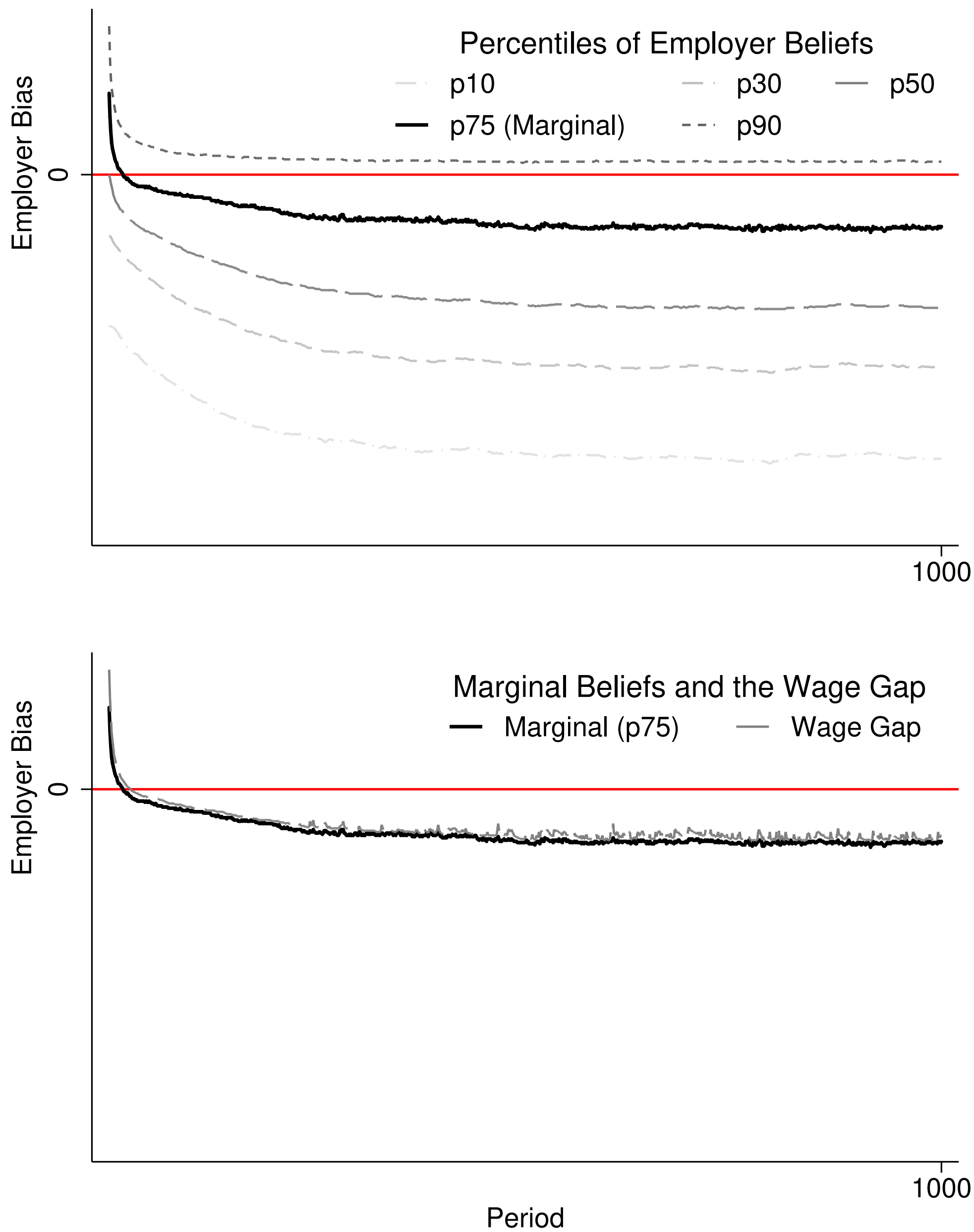
In Figure A2-2, I conduct simulations with entry and exit of employers in which employers below the cutoff are 100% more likely to exit in any given period and new employers enter the market with biased priors. Namely, mean prior beliefs equal the average belief of employers already in the market. A wage gap is sustained and larger than when employers have unbiased priors as in Figure 2, even with a higher differential exit rate.

Figure A2-1: Wage Gap and Model Parameters



Equal Group Size refers to group B being of equal size to group A (50% of workers). Homogenous Priors refers to each employer holding prior $\mu_0 = 0$. Negatively-Biased Priors refers to employers having mean prior beliefs below the true value (-1 vs 0). Low Prior Precision corresponds to a case with prior variance equal to 2. High Signal Precision corresponds to a case with variance in worker productivity equal to 1. Stereotype bias corresponds to a case where employers incorrectly believe group B worker productivity to be 2 when it is 4. Lower Mean Productivity corresponds to a case where mean group B productivity is lower than that of group A (-1 vs 0). Lower Exit Differential refers to a case where employers who hire from group A are 10% more likely to exit the market each period. Lower Aggregate Exit refers to a case where the overall exit rate is 1% each period. See Figure 1 for other parameter choices.

Figure A2-2: Model Simulation with Market Entry and Exit, 100% Exit Differential, Biased Priors



The aggregate exit rate corresponds to 2% each period, with a 100% higher exit rate for employers below the hiring cutoff for group B. New entrants have mean beliefs equal to the mean of employers already in the market. See Figure 1 for other parameter choices.

Appendix 3 - General Productivity Distribution

Let worker productivity be drawn from $X|\mu_B \sim G(x)$, a one-parameter family of distributions characterized by their mean, with full support on an interval of real numbers $\mathbb{X}$, bounded variance, and density function $g(x)$. The parameter of interest is the expected productivity of group B, $\mu_B = E_G[x]$. Employers have a common prior distribution about group B's mean productivity $h(\mu_B)$. Each hire provides an i.i.d. private signal x about worker productivity and S_{jt} is the collection of all signals observed by time t. Under strictly monotone and continuous Bayesian updating on the mean, the distribution of posterior beliefs conditional on S_{jt} corresponds to

$$z(\mu_B|S_{jt}) = \frac{\Pi_{k \epsilon S_{jt}} g_{x_k}(x_k) h(\mu_B)}{\int \Pi_{k \epsilon S_{jt}} g_{x_k}(x_k) h(\mu_B) d\mu_B}.$$

The hiring decision hinges on the expected productivity of Group B, which decreases with lower signals about the group's productivity. As such, hiring decisions, market clearing conditions and wage setting are unchanged, along with Proposition 1. Proposition 2 follows under regularity conditions on $G(\cdot)$ and $h(\cdot)$. Proposition 3 follows from assumptions made on $G(\cdot)$ as well as Propositions 1-2.