Canadian Labour Economics Forum

WORKING PAPER SERIES

Endogenous Learning and the Persistence of Employer Biases in the Labor Market

Louis-Pierre Lepage (U. Michigan)

Conference Volume CLEF – 2020, WP # 24

Endogenous Learning and the Persistence of Employer Biases in the Labor Market

Louis-Pierre Lepage
University of Michigan

April 11, 2020

Abstract

I present a statistical discrimination model of the labor market in which persistent negative employer biases about the productivity of a group of workers arise through hiring and learning about the group. Bayesian profit-maximizing employers endogenously develop biased beliefs based on their hiring experiences which lead to asymmetric learning about the group's productivity across employers. Optimal hiring follows a cutoff rule in posterior beliefs and market-clearing wages below which employers stop hiring from the group, preserving negative biases and leading to a negatively-skewed aggregate distribution of beliefs. Long-run discrimination in the form of a wage below the group's expected productivity can arise even with market competition, without productivity differentials across worker groups or prior employer biases, and regardless of worker signaling or investment decisions. The model generates predictions analogous to the Becker taste-based model, in a statistical framework with beliefs replacing preferences, rationalizing apparent prejudice as the result of "incorrect" statistical discrimination.

* I thank Martha Bailey, John Bound, Charlie Brown, David Miller, Tanya Rosenblat and Jeff Smith for continued guidance and feedback. I thank Aislinn Bohren, Michela Carlana, Kerwin Charles, Ash Craig, Ellora Derenoncourt, Alex Imas, Kevin Lang, Steve Lehrer, Conrad Miller, Nicolas Morales, Mike Mueller-Smith, Mandy Pallais, Rosina Rodriguez Olivera, Mel Stephens, Justin Sydnor, Joanne Tan and Justin Wolfers for helpful discussions as well as seminar participants at the University of Michigan, Labo(u)r day at Michigan State University, the University of Wisconsin-Madison and the EGSC at Washington University in St. Louis for their comments. I am responsible for all errors.

After decades of cultural change and anti-discrimination legislation, African-Americans and women still fare worse than white men in the labor market. One striking feature of these differentials is their persistence. Lang and Lehmann (2012) report that the full-time male black-white earnings ratio remained at 0.77 in 2010 with little progress since the late 1990s, while the unemployment ratio remained constant at 2 between 1968 and 2008. Blau and Kahn (2017) report that the full-time gender earnings ratio has hovered around 0.78 since the late 1990s. The mechanisms through which discrimination may contribute to these disparities remains an open question with implications for theory, empirical work and policy.

Canonical models of statistical discrimination assume that employers have correct equilibrium beliefs about the productivity of worker groups and therefore explain the persistence of discrimination as a rational response to productivity differentials (Phelps, 1972; Arrow, 1973; Aigner and Cain, 1977, Coate and Loury, 1993; Moro and Norman, 2004; Fang and Moro, 2011).[1] These models suggest policies to improve information available on individual workers at hiring and reduce productivity differentials across groups. In contrast, models of taste explain discrimination through exogenous preferences of employers for groups of workers (Becker, 1957; Black, 1995; Charles and Guryan, 2008). They provide an intuitive rationale for differences between average performance and average pay of a group and suggest a different set of policies to mitigate discrimination such as fostering market competition.

This paper presents an alternative statistical discrimination model of the labor market with the key feature that employers learn about group productivity through hiring. When hires provide private information about the group and employers have worse initial information on a particular group's productivity, they trade off learning about that group against

[1]This is necessarily true in models of self-fulfilling prophecies such as Coate and Loury (1993) since workers confirm employer priors through their investments.

current-period profit maximization.[2] An employer's hiring history, therefore, influences their future hiring decisions. Positive experiences create positive biases which correct themselves through more hiring and learning. Negative experiences, however, decrease hiring and learning, resulting in the persistence of negatively-biased beliefs. Asymmetric learning across employers results in a negatively-skewed distribution of beliefs about the group's productivity.[3] Market-clearing wages are determined by the willingness to pay of the marginal employer, which determines optimal hiring. Over time, the distribution of beliefs can cause the wage to fall and remain below the group's expected productivity in the long-run. The model predicts discrimination even with equally-productive worker groups, without prior biases or endogenous worker responses.[4] Biased beliefs arise endogenously from optimal hiring and discrimination can survive competition in the form of higher exit rates for biased employers, contrasting with the view that biased beliefs should be learned or competed away.

This paper contributes to the literature on discrimination and employer learning, studying learning about groups rather than individuals (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007; Arcidiacono et al. 2010; Kahn and Lange, 2014). Even though employers engage in statistical discrimination, my model generates core predictions of taste-based models with beliefs replacing preferences, rationalizing apparent taste-based discrimination as "incorrect" statistical discrimination. Apparent prejudice is the endo-

[2]This setting resembles a bandit problem (Bergemann and Välimäki, 2008). It has similarities with costly information acquisition models, but these models typically consider explicit exogenous costs. Learning costs in my model arise implicitly from inter-temporal profit maximization and are endogenous to an employer's beliefs and those of other employers though market-clearing wages.

[3]Lepage (2020) formally tests this bias-generating mechanism in an experimental labor market and presents evidence supporting its main predictions.

[4]Arrow (1973) mentions that biased employer priors could lead to a self-confirming equilibrium if employers ignore subsequent information or worker responses confirm employer beliefs. I provide a different mechanism through which biased beliefs create discrimination from uncertainty, without deviating from profit-maximization or Bayesian updating. Beliefs in my model are not self-confirmed, but endogenously self-sustained for a fraction of employers through optimal hiring and market clearing.

genous result of experiences shaping beliefs and views in potentially distortionary ways.[5] Similarities between biased beliefs and taste may appear unsurprising at first glance. Yet, the specific way I model biased beliefs highlights that the main insights of prejudice-based models for labor-market discrimination can arise from uncertainty, without deviations from profit-maximization or Bayesian updating.

The model complements the statistical discrimination literature by showing that biased beliefs arise from imperfect information on groups in an otherwise standard setting and that employer learning about some groups can be particularly slow. In the presence of biased beliefs, prejudice and statistical discrimination are not necessarily distinct, consistent with recent evidence (Bohren et al., 2019a; 2019b).[6] The model informs tests aiming to identify the source of discrimination and leads to different policy implications than either statistical discrimination or taste. For example, my model provides a new lens to analyze affirmative action which, through endogenous employer investments, can lead to learning and persistent improved outcomes as documented in Kurtulus (2015) and Miller (2017). Consistent with empirical work reviewed in Lang and Kahn-Lang Spitzer (2020), my model predicts that providing information on groups of workers may mitigate discrimination, as could intergroup interactions (Pettigrew and Tropp, 2006; Paluck et al., 2019).

The paper is organized as follows: Section 1 presents the model while Section 2 presents extensions. Section 3 situates the model in the theoretical literature. Section 4 discusses related evidence as well as implications for empirical tests and policy. Section 5 concludes.

[5] In practice, individuals appear quick to form beliefs about groups and act on these in a way that shapes future views. Beliefs are particularly compatible with context-specific discrimination such as variation across skill and education levels, as is the case empirically (Lang and Lehmann, 2012).

[6] Bertrand and Duflo (2017) mentions that psychology has been more nuanced about this distinction. Prejudice is seen as the result of group membership interacting with social identity or associations from exposure to groups. I show that this nuance is warranted using standard economic models. My model shows 1) how biases can micro-found the reduced-form notion of prejudice in economics and 2) how prejudice in the form of biases endogenously affects decision-making in statistical discrimination models.

1 Labor Market Model

1.1 Employer Information and Beliefs

Consider a setting with a large number of employers and workers from two observably different groups A and B (for example race or gender) who potentially differ in their average productivity due to historical or social factors. I focus on employers from group A, who know the productivity distribution of their own group but are initially uncertain about that of group B.[7] Employers can learn about group B's productivity by hiring group B workers, but their objective is to maximize profits and previous hiring experiences therefore determine both beliefs about the group's productivity and the value of additional learning.

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. Employers have common priors about the mean productivity of group B given by $\mu_B \sim N(\mu_0, 1/\tau_0)$ and I focus on the case where $\mu_0 = \mu$ such that employers start out with 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.[9]

In the baseline model, I make three simplifications relating to hiring and learning. First,

[7] The key feature is that knowledge about the other group is "less certain", but assuming complete information on group A workers makes the analysis and exposition much simpler. Information asymmetry could arise if agents have better information about their own group due to previous experiences and interactions inside or outside the labor market. Similar reasoning has been used in both economics and psychology, but I consider the implications of this information asymmetry for learning about groups (Lang, 1986; Cornell and Welch, 1996; Kelly et al., 2007). Alternatively, information asymmetry could arise in a majority versus minority setting where market participants have better information on the majority.

[8] Appendix 3 generalizes the results to other productivity distributions.

[9] The implications of firm size for the model's predictions are discussed in Section 3. One-period contracts restrict attention to group learning by studying employers repeatedly choosing between the groups. Multi-period contracts may slow down learning if employers retain good workers, but such contracts do not change relative incentives to hire and learn about group B determined by μ_B.

employers observe no individual signal of productivity prior to hiring. They rely solely on group membership to predict the productivity of a worker at the hiring stage. Second, worker signals of productivity are only available through an employer's own hiring. Third, there is no human capital investment or signaling; worker behavior is the same across groups and does not respond to employer beliefs. Each worker is endowed with a fixed productivity and inelastically provides a unit of labor each period. The implications of these simplifications on the model's predictions are discussed in Section 3.

Workers hired from group B determine the information set of employer j, S_{tj}, composed of one private signal drawn from X for each hire. The cumulative number of signals employer j has observed by time t is $K_{tj} = \sum_{n=1}^{t} \mathbb{1}(L_{Bnj} = 1)$ where L_{Bnj} is an indicator variable for whether a group B worker was hired in period n. Employers form posterior beliefs about mean group B productivity according to the Normal updating formula.[10] For employer j at time t,

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

Letting $E[\mu_B|S_{tj}] = \frac{\tau_0\mu_0+\tau\sum_{i=1}^{K_{tj}} x_i}{\tau_0+\tau K_{tj}}$ and $\text{Var}(\mu_B|S_{tj}) = \frac{1}{\tau_0+\tau K_{tj}}$, employers form posterior beliefs about group B productivity $X_B \sim N(E[\mu_B|S_{tj}], \text{Var}(\mu_B|S_{tj}) + 1/\tau)$ which captures the expected productivity of future group B hires.[11]

[10] In practice, learning could be influenced by social categorization, the tendency to think about others in terms of their group membership, which leads to perceived outgroup homogeneity and stereotypes (Allport, 1954). Thinking in terms of group membership could lead employers to overestimate hiring signal precision when updating their beliefs about group B ($\tilde{\tau} > \tau$). Such stereotype formation is not necessary to generate discrimination, but could amplify it as shown in Appendix 2.

[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 $\int \phi_{\mu_B|S_{tj}}(m) \int \phi(x|m)(x - E[\mu_B|S_{tj}])^2 dx dm = \text{Var}(\mu_B|S_{tj}) + 1/\tau$.

1.2 Hiring Decision

Consider a frictionless labor market which clears each period.[12] I postpone the discussion of product-market competition to build intuition with a version of the model with no entry or exit. Employers are wage-takers and maximize the present value of lifetime profits $\Sigma_{t=0}^{\infty}\beta^t E[\pi_{tj}]$.

Employers consider the value of gathering information on the productivity of group B, leading to a dynamic optimization problem. An individual employer's posterior beliefs are characterized by $\psi_{S_{tj}} = \{E[\mu_B|S_{tj}], \text{Var}(\mu_B|S_{tj})\}$ and Ψ_t is a list of posterior beliefs across all employers. Conditional on current beliefs and wages at time t, employer j chooses between hiring from group A or group B to maximize their expected profits

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

Group A's wage, w_A, is time-invariant and equal to their expected productivity μ. Group B's wage, w_{Bt}, evolves under the influence of Ψ_t. The continuation value $V(\cdot)$ includes updated beliefs $\psi'_{S_{t+1,j}}$ when a group B worker is hired and $\psi_{S_{t+1,j}} = \psi_{S_{tj}}$ otherwise. $E_t[V(\psi'_{S_{t+1,j}}, w_{B,t+1}(\Psi_{t+1}))] \geq E_t[V(\psi_{S_{t+1,j}}, w_{B,t+1}(\Psi_{t+1}))]$ since hiring from group B provides information on productivity which cannot decrease expected profits.

For tractability, I assume static wage expectations: employers expect wages to remain constant across periods, $E[w_{B,t+1}] = w_{B,t}$.[13] Static wage expectations imply that employers

[12] Market clearing in the model signifies that demand and supply are equalized for both groups of workers and that each worker-employer pair has no incentive to deviate.

[13] Recent work surveyed in Aguirregabiria and Jeon (2019) increasingly reflects that uncertainty and learning in complex competitive environments may lead firms to have biased beliefs, for example about demand,

cannot recover the aggregate distribution of beliefs from observing the wage and do not form beliefs about the beliefs of other employers. This assumption appears particularly mild in practice given the complexity of the problem faced by employers. Even if they recognize that other employers are Bayesian and could predict the evolution of aggregate beliefs, the relative wage in practice depends on a host of factors (changing skill and education across groups, changes in industry and occupation mixes, demographic changes, etc.) such that separately isolating the impact of employer beliefs on wages seems implausible. As such, taking the current wage as a prediction for the wage next period seems a reasonable approximation in the context of the model, especially since it is correct in the long run.

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

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

The left side of equation (3) represents the expected value of additional information from hiring a group B worker, while the right side represents the expected foregone profit of hiring a group B worker. The perceived value of additional information depends on the likelihood that it will lead to changes in labor allocation and therefore higher expected profits. It is maximized at $\mu_B = \mu$ since information is likeliest to affect future hiring, and decreases as μ_B becomes biased away from μ. In the case of negative bias, hiring group B workers becomes less attractive from both a learning and a production standpoint. Thus, when prior

costs, or the behavior of other firms. Similar assumptions are commonly made in a wide range of dynamic models such as cobweb models, recursively dynamic computable general equilibrium (CGE) models, or the Mundell-Fleming model.

experience suggests that group B is less productive, there is an inherent trade-off between potential future benefits of learning and foregone profits from hiring less productive workers. This trade-off can be represented by a one-arm 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 (Robbins, 1952; Gittins and Jones, 1974). Obtaining comparatively low payoffs from sampling the risky arm eventually leads the agent to stop experimenting and choose the safe arm, with the important difference that wages and therefore payoffs are endogenous in my model.

1.3 Wages and Hiring Cutoff

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

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

The trade-off between learning and foregone profit putting aside wage considerations is captured by λ_{jt} . It can remain above 0 if $E[\mu_B|S_{tj}]$ falls below μ, highlighting that employers may hire workers from group B even if they believe them to be less cost-effective, to avoid potential future losses from incorrect beliefs.

Each period, labor market clearing requires that, given current wages, the fraction of employers who prefer to hire from group B is equal to the fraction of workers from group B. The wage of group B in every 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, $\lambda_{mt} = w_{Bt}(\Psi_t) - w_A$, determining the optimal hiring strategy of employers as 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.

Since the wage gap is determined by λ_{mt}, the optimal hiring decision of other employers immediately follows. Those with λ_{jt} above the marginal hire from group B and others from group A, clearing the market. 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.

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, as 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 $T : \triangle\, \mathbb{R}^2 \rightarrow \triangle\, \mathbb{R}^2$, a wage function $w_B : \triangle\, \mathbb{R}^2 \rightarrow \triangle\, \mathbb{R}$ and an initial state $\Psi_0 \in \triangle\, \mathbb{R}^2$ such that in every period:*

1. *The labor market clears:*

$\Psi_t(\{\psi_{S_t j} : \lambda_{jt} \geq \lambda_t^c(w_{Bt}(\Psi_t))\}) = F_B$ *where F_B is the fraction of workers from group B.*

2. Employers are Bayesian updaters: for any arbitrary set $\Theta \in \mathbb{R}^2$

$$T(\Psi_t)(\Theta) = \Psi_t(\{\psi_{S_{tj}} : \lambda_{jt} < \lambda_t^c(w_{Bt}(\Psi_t))\} \cap \Theta) + \int_{\{\psi_{S_{tj}}:\lambda_{jt} \geq \lambda_t^c(w_{Bt}(\Psi_t))\}} \int_{\mathbb{R}} \{\mathbb{1}(\psi'_{S_{tj}}) \in \Theta | \psi_{S_{tj}}, x\} \phi_{\mu,\tau}(x|\mu) dx d\Psi_t(\psi_{S_{tj}}).$$

3. Employers maximize expected profits according to equation (2) for all $(\psi_{S_{tj}}, w_{Bt})$.

The first 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 second conditions states that employers below the hiring cutoff for group B do not update their beliefs about the group's productivity, while those above the hiring cutoff update their beliefs based on the productivity of their hire according to Bayes' rule.

1.5 Bias and Discrimination

As a result of the optimal hiring rule and equation (1), it is straightforward to characterize the asymptotic distribution of posterior beliefs as 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* $E[\mu_B|S_{tj}] < \mu$*. The limiting fraction of employers with* $E[\mu_B|S_{tj}] < \mu$ *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

$\mu_B < \mu$ in the long run given a strictly positive value of learning) do not hire from group B, preserving negative biases. Similarly, 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.[14] Proposition 2 highlights that optimal hiring and learning lead a subset of employers to hold negatively-biased beliefs, even asymptotically.

The bias-generating mechanism captures a fundamental aspect of hiring in labor markets and generates a plausible distribution of beliefs for discrimination to arise. First, beliefs about group B's productivity exhibit heterogeneity among employers which can be sustained over time. Second, asymmetric learning about the group's productivity across employers results in aggregate beliefs being negatively-skewed. The mechanism generates these features without relying on group differentials (ex-ante or ex-post)[15] or bias as a primitive of the model, providing a novel way to understand persistent, heterogeneous, asymmetric beliefs about a group of workers in a labor-market setting. The mechanism complements work on how bias may affect the belief updating process itself (Sarsons, 2017) or the way workers are evaluated and supervised (Bartoš et al., 2016; Glover et al., 2017), but highlights that such behavior is not necessary for biased beliefs to create labor market discrimination.

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

[14]The Becker taste-based model requires that the fraction of prejudiced employers be at least as large as the fraction of group A workers. 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 Lang and Lehmann (2012) discusses evidence that a large share of employers hold negative perceptions in the context of black workers. Moreover, Black (1995) shows that wage gaps may be sustained under milder conditions in a search framework, as discussed in Section 3.

[15]This distinction has important implications even if it is unlikely that both groups have equal productivity in practice, since it predicts that policies aiming at closing productivity gaps between groups would not necessarily eliminate discrimination if information asymmetries remain. See Bordalo et al. (2016) for a model of stereotypes in which agents distort true group differentials.

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 is stable. With hiring experience, supramarginal values of λ_{jt} become concentrated around 0 as $E[\mu_B|S_{tj}]$ 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 cutoff are expected to have negative hiring experiences with group B in any given period, such that their λ_{jt} fall below that period's cutoff. If the wage does not decrease, then the market does not clear. Lastly, since beliefs are fixed asymptotically, in the long run there is virtually no updating and no change in the wage, so it converges to a constant. Given the assumption that both groups of workers are equally productive, a wage gap also implies that group B is paid below their expected productivity. While the predicted wage gap across groups depends on their relative productivity, the result that group B will be paid below their expected productivity doesn't.

The model thus predicts that biased beliefs systematically arise from endogenous learning about group B, remain in the long run, and generate persistent discrimination against group B. The wage of group B falls and remains below its mean productivity regardless of true productivity differentials with group A or employer priors.

1.6 Entry, Exit and Competition

A common view in the literature is that market competition should drive out biased beliefs and therefore discrimination, at least in the long run. I thus turn to a version of the model in which some employers enter and exit the market each period. Entry and exit do not change the fundamental intuition of the bias-generating mechanism regarding asymmetric learning and biased beliefs, but provide 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 aggregate rate δ every period. Hiring decisions and wage determination follow the same process as before. Profit maximization is given by

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

Three additional elements influence the expected wage gap: the exit rate, how the exit rate depends on employer beliefs, and priors held by new employers. The exit rate itself governs the expected length of survival in the market, incentives to learn about group B, and time available for employers to potentially correct their biases. It can also directly affect the aggregate distribution of beliefs by replacing experienced employers with new employers who may hold different beliefs on average. This is especially true if the exit rate is higher for employers with negatively-biased beliefs (δ_{Biased}), capturing the notion of competition.[16] I consider two benchmark cases for priors held by new employers, which affect the aggregate distribution of beliefs and the value of learning for new entrants. First, I consider the case in which new employers always enter the market with unbiased priors. Second, I consider the

[16] Employers know the aggregate exit rate, but not whether they face an increased exit rate specifically due to their beliefs. Similarly to static wage expectations, the exit rate of an employer in practice is a complicated function of a variety of factors, such that it may be infeasible for an employer to isolate the impact of their beliefs about group B productivity on their exit probability.

alternative that new employers have their initial beliefs influenced by experienced employers. Specifically, I assume that the mean prior beliefs of new employers is equal to the average posterior beliefs of employers already in the market. This assumption reflects the idea that new employers believe that experienced employers hold correct beliefs on average, as assumed in much of the literature, but may lead new entrants to hold biased priors.[17]

Across these considerations, the major insight is that bias arises endogenously from optimal hiring. Therefore, as some employers held unbiased priors but developed biased beliefs through hiring, so may new employers. In the aggregate, biased beliefs do not generally go away with competition. As a result, competition may disproportionately drive out biased employers, but the wage gap will not necessarily be eliminated.[18] Depending on the parameters governing entry and exit, a wage gap can be sustained asymptotically even if biased employers are driven out at a higher rate, as summarized in Remark 1.

Remark 1 (Persistent Discrimination with Market Competition)
For some values of δ and δ_{Biased}, 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 from Proposition 3 is that the existence of a wage gap depends on model parameters because

[17] Another consideration is whether prior variance of new employers may decrease if they learn from previous "generations" of employers. If so, the impact of negative experiences may be mitigated over time. This is unlikely to eliminate the initial information asymmetry since it would require employers to completely ignore their own hiring experiences. Further, the learning problem faced by employers in practice is constantly changing such that they must rely on their own experiences to assess group productivity in their hiring context. For example, the relative productivity of women and minority workers compared to that of white men was not the same in 1980 as it is today, and employment contexts have changed substantially.

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

competition can affect hiring and learning as well as the distribution of employer beliefs in the market. For exit rates near zero, the existence of a wage gap directly follows from Proposition 3. For extremely 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, a higher exit rate or higher differential exit rate for biased employers reduces the magnitude of the wage gap, as shown in appendix 2 and consistent with empirical evidence such as Ashenfelter and Hannan (1986) and Black and Strahan (2001). At the extensive margin, the model predicts that competition may not eliminate discrimination arising from biased beliefs, even in the long run.

1.7 Simulations

To illustrate the model's dynamics, a set of simulations was computed with 10,000 employers and 10,000 workers, 25% of which are from group B.[19] Worker productivity is distributed $N(0, 2)$ and prior beliefs are distributed $N(0, 1)$. w_A is normalized to 0 and β is set to 0.9. Because the simulated market is finite, the evolution of beliefs and wages is stochastic rather than deterministic. Emphasis from the simulations should be put on the general dynamics of the model characterized by Propositions 1-3 and Remark 1, which do no substantively vary with parameter choice, rather than specific values of the wage gap or the speed with which wages evolve. In particular, the existence of a negatively-skewed distribution of beliefs and a wage gap do not depend on parameter choice without entry and exit. With entry and exit, whether biased beliefs lead to a wage gap depends on parameter choices. Similarly, the

[19]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.

initial state in which employers enter the market at the same time and hire workers for the first time exhibits theoretically intuitive features, but is of limited practical interest.[20] See Appendix 2 for simulation results showing how the wage gap changes with the model's main parameters.

Panel A of Figure 1 shows the evolution of beliefs for key moments of the distribution over 1,000 periods without entry and exit. The 25% of employers with the highest valuation for group B workers each period are those above the hiring cutoff who hire them and learn, such that their beliefs become distributed around the true value while those of other employers lie below. Beliefs of employers above the cutoff converge towards zero (76th percentile and above in Figure 1), while those of other employers do not evolve. Negative biases are sustained, and it may in fact be optimal for employers with low prior beliefs to never hire from group B.

Panel B of Figure 1 shows the group B wage along with the beliefs of the marginal employer. The wage is initially above the beliefs of the marginal employer due to the value of learning about group B. It eventually falls below zero and approaches the marginal employer's beliefs, such that group B starts earning a lower wage as the marginal employer's beliefs fall below μ and the value of learning falls. The wage of group B remains below that of group A in every subsequent period. With a finite number of employers, a separation in the beliefs and willingness to pay of employers above and below the cutoff is created as seen in Panel A contrasting the 75th and 76th percentiles. The relative WTP of employers above converges to zero, while that of employers below remains constant below zero. The

[20] Given unbiased priors and the value of information, market clearing can technically require that the initial group B wage be higher than that of group A. This initial condition is of limited interest in practice since it represents a scenario where all employers simultaneously enter a new market with no hiring experience, and can be overturned or diminished under a wide range of assumptions regarding prior beliefs, relative uncertainty and productivity across groups, or ambiguity aversion.

market clearing wage could technically lie anywhere between the relative WTP of the last employer to hire from group B (76th percentile) and the first employer to hire from group A (75th percentile), while the latter determines the wage with a continuum of employers as characterized in Proposition 3. If workers have no bargaining power and match surplus is allocated to employers, then the wage is also set by the first employer to hire from group A with a finite number of employers as shown in Panel B.

Figures 2 and 3 present simulation results with entry and exit of employers over 1,000 periods. They correspond to a 2% aggregate exit rate each period with biased employers 25% and 100% more likely to exit. In Figure 2, new employers hold unbiased priors. In Figure 3, new employers enter the market with mean prior beliefs equal to the average of employers already in the market. In both cases, a wage gap is sustained and it is larger when new employers eventually hold biased priors as in Figure 3, even with a higher differential exit rate. The set of employers in the market is expected to have been jointly replaced 3 to 4 times over the period shown, so the pattern is simply repeated beyond these periods.

2 Model Extensions

2.1 Outside Learning

Thus far, employers observe information about group B only through their own hiring. If employers observe outside signals, such as the hiring behavior of a competitor or the performance of group B workers in other settings, these signals could affect beliefs and learning in the absence of hiring.[21] Outside learning can either mitigate or exacerbate bias,

[21] For example, outside learning could arise from manager or employer turnover leading to the sharing of knowledge and experiences from different hiring contexts.

but does not change the core predictions of the model with entry and exit.

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

$$\mu_B | S_{tj} \sim N\left(\frac{\tau_0\mu_0 + \tau\sum_{i=1}^{K_{tj}} x_i + \tau_o \sum_{m=1}^{t} o_m}{\tau_0 + \tau K_{tj} + \tau_o t}, \frac{1}{\tau_0 + \tau K_{tj} + \tau_o t}\right). \tag{4}$$

If employers weight outside signals the same as their own experiences, then they effectively observe the result of one "free" hire each period. If $\tau_o < \tau$, employers put more weight on their first-hand experiences, which is likely to arise if there is mismatch between employment contexts. The weight given to these signals determines the extent to which they influence updating, but as long as they are assigned nonzero weight, employers will learn from them.

Employers who hire from group B still learn faster if they observe both hiring and outside signals. Then, the distribution of beliefs remains negatively-skewed in any finite period, and the bias-generating mechanism can be seen as slowing down learning rather than stopping it. Conceptually, with entry and exit, the existence of a long-run wage gap depends on the parameters governing the relative speed with which employers become biased, correct their bias through hiring or outside sources, or exit the market.

Slowing down learning itself has non-negligible implications for discrimination. Statistical discrimination models generally predict that the market immediately learns the productivity of worker types in equilibrium, influencing investment and signaling decisions of workers. One criticism is that learning is "too fast" for these models to be important in the long run (Lang and Lehmann, 2012). As such, my model provides a justification for why learning about some groups of workers can be particularly slow.

Further, the effect of outside signals is not unambiguously positive. First, they lower incentives for employers to hire from group B and learn from their own signals, potentially leading to free-riding.[22] Equation (4) also assumes that outside signals are unbiased, unambiguous and unrelated to existing bias. Otherwise, outside signals could exacerbate discrimination. This is particularly relevant since the majority of employers in the model may hold biased beliefs, potentially leading to the spreading of bias depending on how employers share information.[23]

Outside learning suggests that discrimination may differ across settings where the observability of competitors, workers, and output vary. Similarly, there may be a role for the provision of information if the acquisition of outside signals is endogeneized.

2.2 Signals of Individual Productivity

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

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

[22]See Keller et al. (2005) for bandit problems with multiple players. Multiple players leads to a dynamic public-good problem with free-riding, consistent with evidence from Hoelzemann and Klein (2018).

[23]See for example DeGroot (1974). Also relevant is work related to endogenous social networks and media sources as well as ambiguity of signals and non-Bayesian updating (Gentzkow and Shapiro, 2006; Baliga et al., 2013; Enke and Zimmermann, 2017; Fryer et al., 2018).

where $\gamma = \frac{1/\tau}{1/\tau + 1/\tau_\varepsilon}$ is a measure of the signal's precision. Negatively-biased beliefs about the mean productivity of group B can arise just as before, but the question is whether individual workers can overcome this bias by signaling that they are not representative of the mean worker of their group. This signaling is of little consequence in the model due to sorting, because employers above the hiring cutoff are willing to pay more for a group B worker conditional on a given signal value. Accordingly, hiring and learning dynamics are unchanged, along with resulting discrimination. Workers in the model 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.

With individual signals of productivity, bias and discrimination may vary across occupation, skill, and education levels 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.[24] An important aspect of discrimination empirically is that it 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, the variability 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).

2.3 Endogenous Worker Investments

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 pro-

[24] This suggests that employee testing at the hiring stage could be used to mitigate the learning problem by gathering information on individuals.

ductivity 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.

Contrastingly, my model predicts bias and discrimination even when worker investments are exogenous. Rather than assuming bias, the model motivates the existence of biased beliefs in an uncertainty setting and explicitly models their evolution. In contrast to models of "self-fulfilling" prophecies such as Coate and Loury (1993), employers in my model hold heterogeneous beliefs about group B. 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 accumulation 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 workers earn lower returns from the labor market overall, they may have incentives to invest less in human capital in general.

2.4 Firm Size

Firm or establishment size has natural implications for learning and biased beliefs. Employers who hire more workers have a higher value of learning and may learn more quickly. As a result, negative biases may be less likely to arise and persist, and these employers may hire a higher fraction of group B workers on average. It is not clear whether a higher fraction of group B workers is a cause or a consequence of size, but it is 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.[25]

Implications for the wage gap are limited assuming that each establishment hires a negligible fraction of the market and that there is sufficient size heterogeneity above the hiring cutoff. Unless all of group B is hired by large establishments with centralized hiring, then these establishments will not be marginal, by definition, and the wage gap will be determined by smaller establishments. In practice, 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. For example, 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.[26]

2.5 Search Frictions

While a formal search model is beyond the scope of this paper, the discrimination literature suggests that search frictions may have important implications. In a random matching setting, the intuition behind the bias-generating mechanism is unchanged. Employers who hire from group B and have negative (positive) experiences are less (more) likely to select a worker from the group again in the future. Positive biases are learned away more quickly than negative ones, so beliefs are negatively-skewed. Nevertheless, the wage gap in these

[25]Evidence from hiring at decentralized firms suggests that 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).

[26]This is based on Headd (2000) which provides the proportion of black workers across firm size in 1998 combined with statistics from the Census Bureau on the total number of workers employed at firms below 25 workers and the total number of black workers for the same year.

models is determined by the average rather than the marginal employer, as highlighted by Black (1995) in reference to the Becker model. Accordingly, wage gaps may be larger and more prevalent in a search framework.[27]

2.6 Minority Employers

The impact of minority employers depends on whether they share the beliefs of the majority or face a different learning problem (they know group B productivity but must learn about group A). In the first case, the distinction between employer types is irrelevant for purposes of bias and wages. In the second case, these employers constitute a fraction of the market who may not develop biases about group B. If their share is large enough, then this encourages segregation and may mitigate wage gaps. Other factors may make it difficult for minority employers to be successful. They may face uncertainty about the majority group, face similar types of discrimination in promotion or the capital market, or make lower human capital investments from lower labor market returns more broadly. Further, empirical evidence for both race and gender suggests that the proportion of managers is relatively low compared to that of workers (Giuliano et al., 2009; Blau and Kahn, 2017).

3 Relationship with Other Theories of Discrimination

My model relates intuitively to theories of taste-based and statistical discrimination. It is a substitute for some of the role of taste-based discrimination and a complement to previous models of statistical discrimination, while remaining substantively different from both.

Biased beliefs have been a controversial concept in the literature since they are often

[27]Search frictions could also mitigate the stark prediction that employers below the hiring cutoff in the long run never hire from group B again.

seen as arising from employer mistakes. Aigner and Cain (1977) state in their influential statistical discrimination 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." This argument captures the prevailing view in economics that biases should not persist if they arise at all, for example, because of learning and competition. It serves to justify the assumption that employers know the productivity of groups made in more recent statistical discrimination work (Coate and Loury, 1993; Cornell and Welch, 1996; Knowles et al., 2001; Moro and Norman, 2004; Fryer, 2007; Morgan and Vardy, 2009) and why little work considers biased beliefs in a labor market setting. Further, if biased beliefs are viewed as inefficient mistakes arising from cognitive biases, they may intuitively look similar to taste-based discrimination in many ways, and it's unclear whether differentiating between the two is meaningful in many contexts.

The model shows how misleading this line of thinking can be. Modeling employer learning about the productivity of worker groups recasts biased beliefs as the systematic outcome of profit-maximization by Bayesian employers. It shows that biased beliefs can play a role in determining equilibrium discrimination in economic markets and have different implications than existing models of statistical discrimination or taste.

First, relating to taste-based discrimination, my model generates predictions analogous to those from Becker (1957) with preferences replaced by beliefs and set in a statistical discrimination framework:

1. An employer prefers to hire workers of group A if the wage gap is smaller than λ_{jt}

and group B otherwise.

2. If enough employers have (approximately) correct beliefs, group B will be hired by these employers. This will effectively result in segregation but no wage gap.

3. If the share of employers with biased beliefs is large enough, there will be a wage gap determined by the marginal employer.

The model has many of the same intuitive implications as taste-based discrimination, namely a difference between average productivity and average pay, but without deviating from profit-maximization. This is a key point given that taste-based discrimination has often been criticized for the arbitrariness of including prejudice in the utility function. My model highlights that the important insights of taste-based models for labor market discrimination do not rely on preferences, but can be understood as arising from information uncertainty.

Biased beliefs capture nuanced aspects of discrimination which are less compatible with the classical notion of taste as an aversion to contact. In many cases, apparent prejudice seems context dependent.[28] Models of taste must also rely on a high fraction of prejudiced employers to generate discrimination (Lang and Lehmann, 2012), and the notion of widespread biased beliefs may be more plausible, especially given work such as Bertand et al. (2005) on implicit discrimination.

This does not imply that taste and biased beliefs are incompatible or cannot co-exist. In fact, they could interact in important ways, and if biased beliefs are reinforced through behavioral primitives in the utility function, become essentially indistinguishable.[29] The two theories may seem difficult to distinguish, but differ fundamentally in how discrimination ari-

[28] For example, it may seem odd to assert that a large share of men dislike interacting with women. It seems more compelling that they may develop biased beliefs about them in some contexts which in turn influence their behavior towards them.

[29] Individuals with a taste for discrimination may be inclined to gather and interpret information in a way that validates and justifies their prejudice. See Nickerson (1998) for an overview of confirmation bias.

ses, evolves, and can be mitigated. One key is the role of information, which has predictable effects on beliefs and is closely related to concepts of statistical discrimination.

Biased beliefs and statistical discrimination are distinct but complementary. In many contexts, the assumption that individuals have complete information on the distribution of characteristics of interest in other groups seems implausible. Discrimination caused by biased beliefs can arise in settings where there are absolutely no grounds for statistical discrimination. It does not arise from employers using *objective* information about groups but using their potentially flawed beliefs to inform behavior. It is not a self-confirming equilibrium 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. My model can sustain discrimination without prior bias and under employer heterogeneity, which is typically ruled out in statistical discrimination. The homogenous-prior assumption made in those models is seldom discussed but crucial to generate long-run discrimination.[30] Yet, my model makes clear that it is not obvious how one could write a model of belief formation and employer learning which would justify this homogeneity assumption, even in the long run.

Another point concerns efficiency and equality. In statistical discrimination models, equilibrium outcomes reflect true average productivity and employers correctly use information at the hiring stage to make rational, profit maximizing decisions. Each group's mean wage is equal to its mean productivity, and ending this type of discrimination under perfect information may not help group B on average. As a result, this type of discrimination is generally regarded as efficient, setting aside notions of "bad equilibria" from models of self-fulfilling

[30] Otherwise, some employers may be relatively better at interpreting signals from the group discriminated against (Aigner and Cain, 1977) or have more accurate beliefs about the group (Coate and Loury, 1993). Other employers would learn or be driven out of the market, such that the need for the discriminated-against group to adjust is unclear.

prophecies. In my model, employers efficiently maximize profits, but workers are paid below their marginal product 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 or welfare cost through increased employer learning.

4 Implications for Empirical Work and Policy

Biased beliefs about groups of workers are consistent with a growing body of evidence and can explain discrepancies between average productivity and pay based on information.[31] They have important implications for identifying the cause of discrimination, which has traditionally 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. That is, if a group appears 10% less productive than another, are outcome differentials consistent with such a productivity gap? If so, this discrimination is classified as statistical and the residual as taste. Such logic is conceptually inadequate to identify the cause of discrimination. As the model makes clear, concluding that the absence of observable productivity differentials implies a taste for discrimination is unwarranted. Residual discrimination could simply arise from attempts at statistical discrimination distorted by incorrect beliefs. Similarly, documenting responses of discriminators to information about individuals and their productivity may be consistent with statistical discrimination, but does not imply that discriminators hold correct beliefs on average or use information correctly.

Recent work such as Arnold et al. (2017) and Bohren et al. (2019b) explicitly considers

[31]See for example Fershtman and Gneezy, 2001; Bertrand et al., 2005; Wolfers, 2006; Reuben et al., 2014; Bordalo et al., 2016; Mobius et al., 2016; Bertrand and Duflo, 2017; Laouénan and Rathelot, 2017; Van Dalen and Henkens, 2017; Sarsons, 2017; Arnold et al., 2018; Dianat et al., 2018; Landsman, 2018; Lesner, 2018; Bohren et al., 2019a; 2019b; Bordalo et al. 2019.

beliefs in their empirical analyses and provides evidence of their importance. Bohren et al. (2019a) make clear from their review that this distinction is rarely taken into consideration in the discrimination literature. They present an example from a lab experiment that tests aiming to identify the cause of discrimination can be misleading if biased beliefs are ignored; and that biased beliefs can be confounded with taste-based or statistical discrimination. My model fills a gap by providing a new bias-generating mechanism, showing that the mechanism generates biased beliefs which affect long-run labor market outcomes, and formalizing the relationship between theories of discrimination. It justifies studying biased beliefs and interpreting evidence of discrimination as arising from biased beliefs even in competitive markets. Therefore, the model highlights that biased beliefs should not be ignored as a potential source of labor market discrimination.

Distinguishing between sources of discrimination is important in part because policy implications can be very different. Increasing competition can mitigate discrimination in the case of taste or biased beliefs, but may not eliminate it even in the long run if information asymmetries remain.[32] Efforts to close productivity gaps may mitigate discrimination if it is based on true group differentials, but my model makes clear that such differentials are not necessary for discrimination to persist. Diversity training and other such interventions may provide relevant information about groups, but if they mainly target cognitive biases and implicit stereotypes, may not address biased beliefs as in my model. Under both correct and incorrect statistical discrimination, providing better information on individual productivity may mitigate discrimination by decreasing employers' reliance on group membership, but information about groups provides an interesting avenue to distinguish between the two.

[32]Following the famous criticism of the Becker model regarding temporariness, some extensions have been proposed by Black (1995), Lang et al. (2005) and Charles and Guryan (2008) in which prejudice may remain in the market under specific assumptions related to imperfect information, adjustment costs, or if prejudiced employers are also prejudiced consumers and coworkers.

An information shock about group productivity may have little impact if firms hold correct beliefs (at least on average), but may mitigate discrimination if it leads firms to update their beliefs.

Indeed, central to the model is the idea that employers can learn about groups through interaction, consistent with recent empirical evidence surveyed in Lang and Kahn-Lang Spitzer (2020). Accordingly, my model has implications for worker subsidies and affirmative action, which can push employers to hire more workers from group B and learn.[33] For example, Kurtulus (2015) and Miller (2017) find that firms under Affirmative Action contracts kept increasing their hiring of black workers even after the contracts ended. One interpretation of this finding is that these contracts induced increased interactions with black workers, leading employers to update their beliefs about the group and hire more of them even after the policy ended. Relatedly, Finseraas et al. (2016) and Dahl et al. (2018) provide evidence in the context of the Norwegian military that simply giving information about women soldiers does not decrease discrimination from male soldiers in the evaluation of candidates, but that intense collaborative exposure and integration does. Pettigrew and Tropp (2006) and Paluck et al. (2019) conclude from their review of the literature that contact typically reduces prejudice. One historical example is World War II which is often discussed as a shock through which firms learned about the productivity of women and minorities, particularly in manufacturing settings (Goldin, 1991). Similarly, Leung (2017) finds that negative and positive hiring experiences of workers from particular countries on an online job board affect the subsequent likelihood of hiring workers from those countries.

Lastly, the model has implications regarding optimal HR policy. Firms may find it opti-

[33] As in Coate and Loury (1993), affirmative action policies could also "backfire" through endogenous worker investments if they push group B to under-invest. Rather than focusing on workers, my model provides a new lens to study these policies by considering endogenous investments by employers.

mal to engage in statistical discrimination, but this can lead to biased beliefs based on hiring interactions which lower expected profits. Then, one distinction with other forms of discrimination is that firms have incentives to eliminate discrimination based on biased beliefs, particularly those with decentralized hiring in which individual managers hold discretionary power. Many considerations go into designing a HR system, but audit study evidence from Berson et al. (2019) suggests that discrimination among large firms appears lower at firms with centralized hiring.

5 Conclusion

This paper presents a statistical discrimination model in which biased beliefs about the productivity of a group of workers create persistent labor market discrimination in the form of a wage below the group's expected productivity. The model shows how biased beliefs can systematically arise through endogenous learning about the productivity of groups based on an employer's hiring history. It then shows that these biased beliefs can create persistent discrimination in labor markets with profit-maximizing, Bayesian employers operating in competitive markets with equally-productive groups of workers, no prior bias, and without endogenous worker signaling or investment decisions.

The model generates the main predictions of the Becker model in the absence of taste for discrimination, highlighting that some of what is generally classified as taste may be better understood as biased beliefs. It can be seen as a substitute for some aspects of taste-based discrimination and provides a new way to understand prejudice in labor markets. Yet, it remains an information story. The model is a statistical discrimination model in which the assumption that employers hold correct beliefs about the productivity of groups of workers is relaxed and beliefs are modeled explicitly. Accordingly, it closely relates to other concepts of statistical discrimination and complements them. Biased beliefs have key implications for understanding the relationship between theories of discrimination, empirical work aiming to identify the source of discrimination, and policies aiming to mitigate it.

The model focuses on profit-maximizing Bayesian employers, but existing work documents a range of behavioral elements which could influence and amplify discrimination based on biased beliefs. As such, the effect of biased beliefs on discrimination as considered in my model may constitute a lower bound in many settings. A natural next step is to explicitly relate biased beliefs as studied in this paper to these behavioral elements, move away from purely mistaken beliefs, and push the relationship with prejudice further. Lastly, the bias-generating mechanism proposed in this paper can be seen as much more general than labor market discrimination and could have important implications in other contexts where agents learn based on their previous experiences.

References

[1] Aguirregabiria, V. and Jeon, J., 2019. Firms beliefs and learning: Models, identification, and empirical evidence. Review of Industrial Organization, pp.1-33.

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

[3] Allport, G.W., 1954. The nature of prejudice. Oxford, England: Addison-Wesley

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

[5] 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.

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

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

[8] 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.

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

[10] 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.

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

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

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

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

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

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

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

[18] 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.

[19] 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.

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

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

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

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

[24] Charles, K.K. and Guryan, J., 2008. Prejudice and wages: an empirical assessment of Beckers The Economics of Discrimination. Journal of Political Economy, 116(5), pp.773-809.

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

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

[27] Dahl, G., Kotsadam, A. and Rooth, D.O., 2018. Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams (No. w24351). National Bureau of Economic Research.

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

[29] DellaVigna, S. and Paserman, M.D., 2005. Job search and impatience. Journal of Labor Economics, 23(3), pp.527-588.

[30] Dianat, A., Echenique, F. and Yariv, L., 2018. Statistical discrimination and affirmative action in the lab. CEPR Discussion Paper No. DP12915

[31] Doob, J., 1949. Application of the theory of martingales. Actes du Colloque International Le Calcul des Probabilités et ses applications, Paris CNRS, 23-27.

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

[33] Fang, H. and Moro, A., 2011. Theories of statistical discrimination and affirmative action: A survey. In Handbook of Social Economics (Vol. 1, pp. 133-200). North-Holland.

[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] Finseraas, H., Johnsen, .A., Kotsadam, A. and Torsvik, G., 2016. Exposure to female colleagues breaks the glass ceiling - Evidence from a combined vignette and field experiment. European Economic Review, 90, pp.363-374.

[37] Fryer Jr, R.G., 2007. Belief flipping in a dynamic model of statistical discrimination. Journal of Public Economics, 91(5-6), pp.1151-1166.

[38] 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, 0(0), pp.1-32.

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

[40] Gittins, J. and Jones, D., 1974. A dynamic allocation index for the sequential design of experiments. Progress in statistics, pp.241-266.

[41] 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.

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

[43] 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.

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

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

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

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

[48] 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.

[49] Karabarbounis, L. and Neiman, B., 2013. The global decline of the labor share. The Quarterly Journal of Economics, 129(1), pp.61-103.

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

[51] Kelly, D.J., Quinn, P.C., Slater, A.M., Lee, K., Ge, L. and Pascalis, O., 2007. The other-race effect develops during infancy: Evidence of perceptual narrowing. Psychological Science, 18(12), pp.1084-1089.

[52] Knowles, J., Persico, N. and Todd, P., 2001. Racial bias in motor vehicle searches: Theory and evidence. Journal of Political Economy, 109(1), pp.203-229.

[53] Kurtulus, F.A., 2015. The impact of affirmative action on the employment of minorities and women over three decades: 1973-2003. Upjohn Institute Working Paper No. 15-221.

[54] Landsman, R., 2018. Gender differences in executive Departure. Unpublished. Bucknell University.

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

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

[57] 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.

[58] Lang, K., Manove, M. and Dickens, W.T., 2005. Racial discrimination in labor markets with posted wage offers. American Economic Review, 95(4), pp.1327-1340.

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

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

[61] Lepage, L., 2020. Endogenous Belief Formation in Hiring and Discrimination. Unpublished. University of Michigan.

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

[63] 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.

[64] 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.

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

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

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

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

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

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

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

[72] 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.

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

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

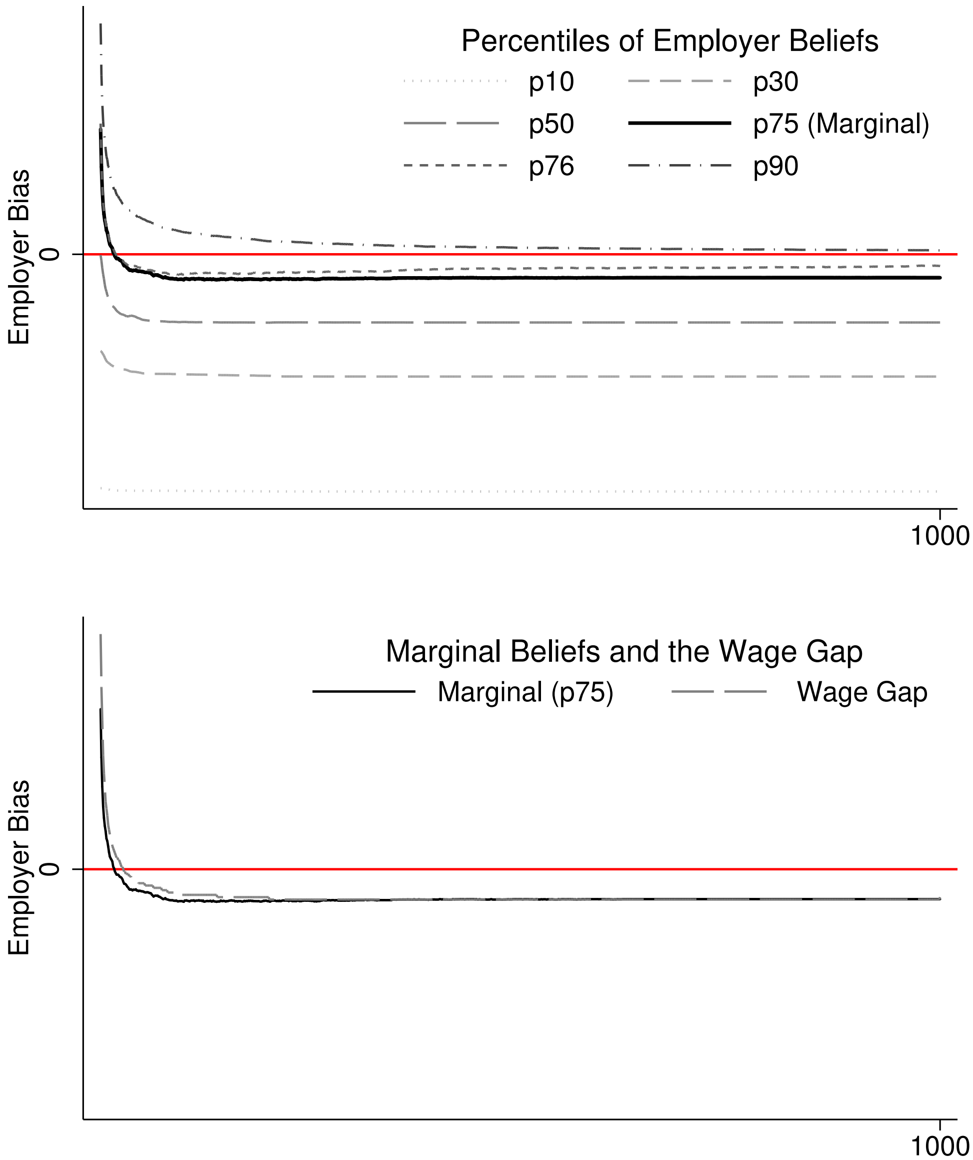
[75] Robbins, H., 1952. Some aspects of the sequential design of experiments. Bulletin of the American Mathematical Society, 58(5), pp.527-535.

[76] Sarsons, H., 2017. Interpreting signals in the labor market: Evidence from medical referrals. Unpublished. University of Toronto.

[77] Van Dalen, H.P. and Henkens, K., 2017. Do stereotypes about older workers change? Evidence from a panel study among employers. CentER Discussion Paper Series No. 2017-028.

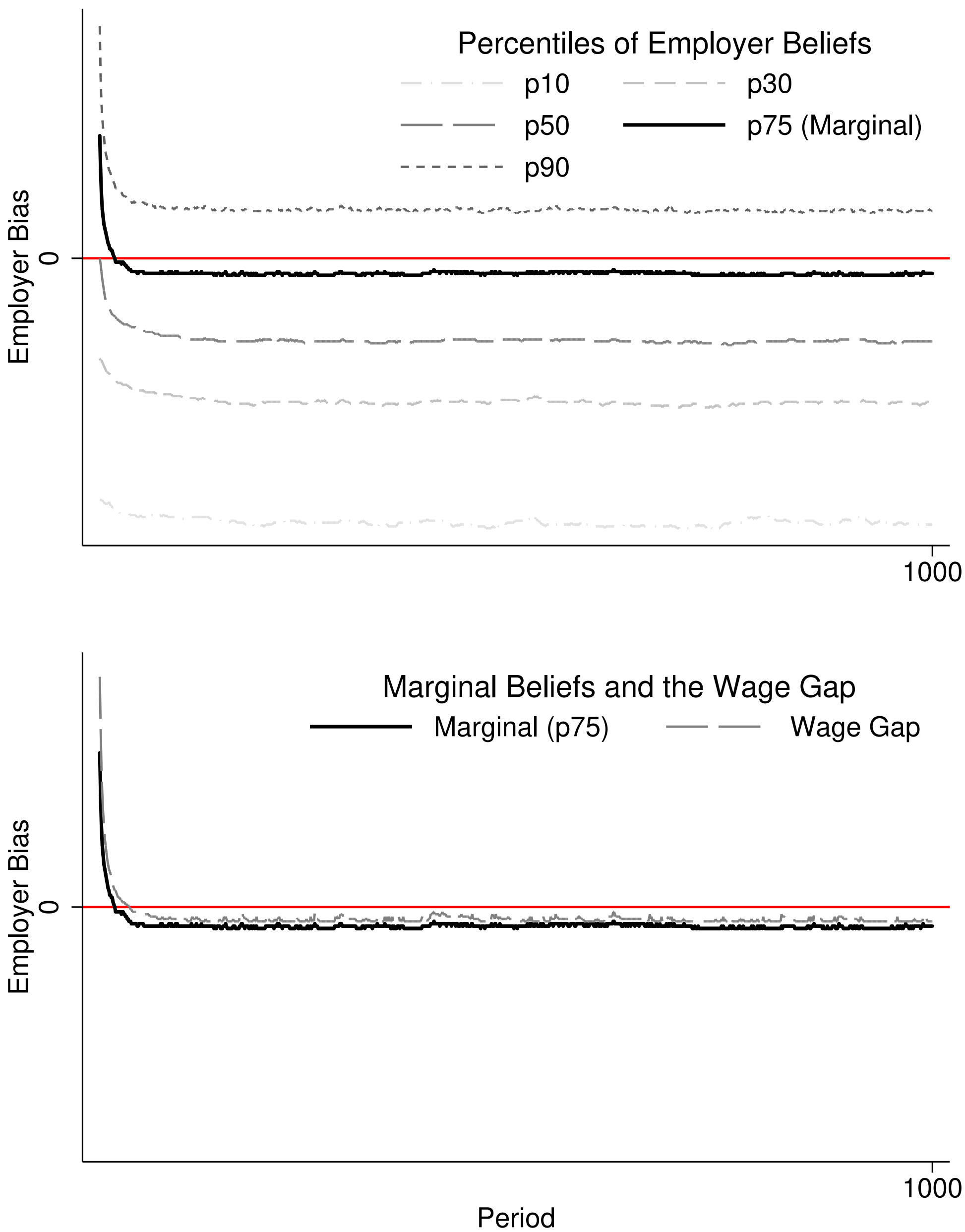
[78] Wolfers, J., 2006. Diagnosing discrimination: Stock returns and CEO gender. Journal of the European Economic Association, 4(23), pp.531-541.

Figure 1: Model Simulation without Entry and Exit



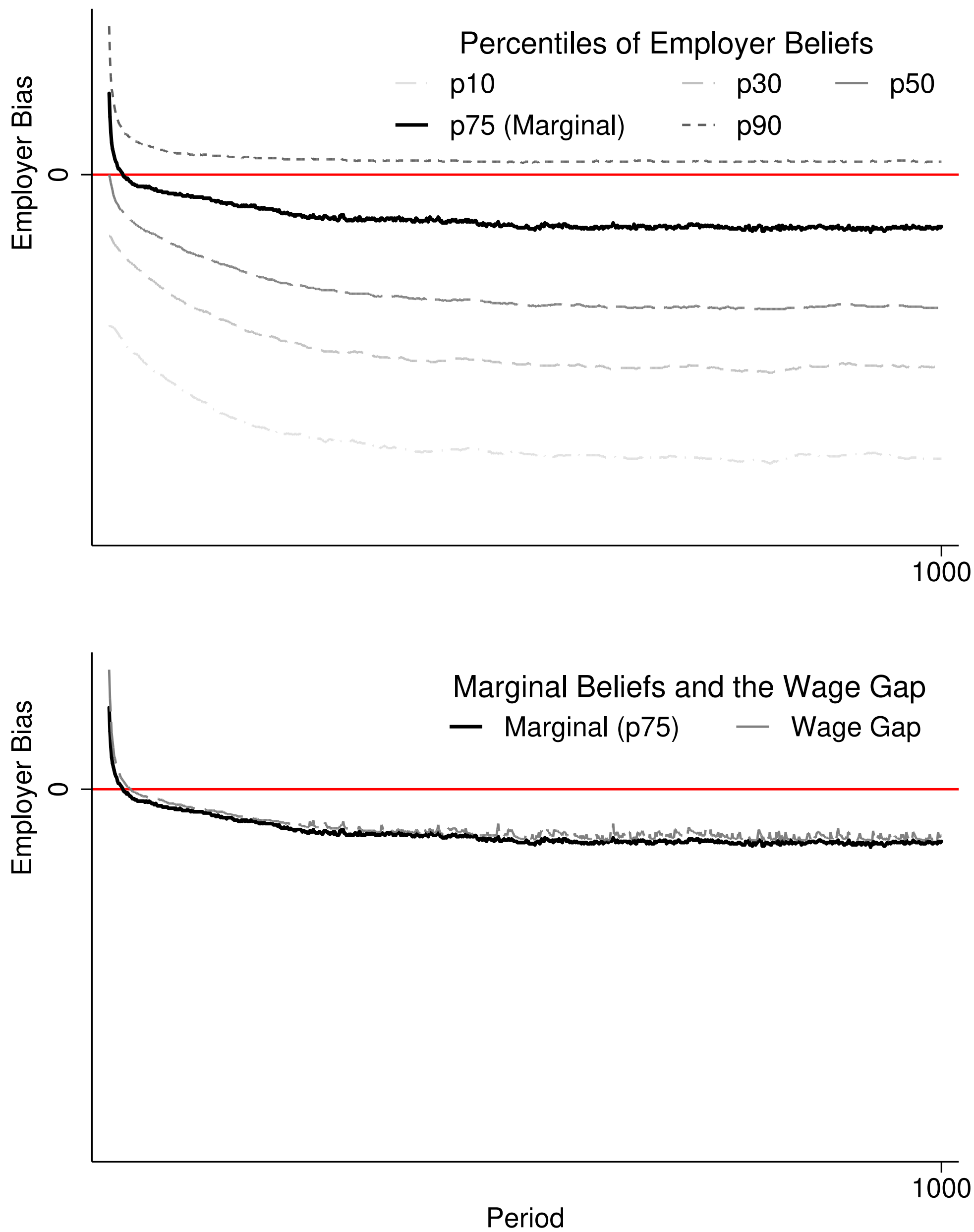
The fraction of group B workers is 0.25. Worker productivity is distributed $N(0, 2)$, prior beliefs are distributed $N(0, 1)$. w_A is normalized to 0 and β is set to 0.9.

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



The aggregate exit rate corresponds to 0.02 each period, with a 100% higher exit rate for employers with negatively-biased beliefs. New entrants have mean beliefs equal to 0 (unbiased). See Figure 1 for other parameter choices.

Figure 3: Model Simulation with Market Entry and Exit, 100% Exit Differential, Biased Priors



The aggregate exit rate corresponds to 0.02 each period, with a 100% higher exit rate for employers with negatively-biased beliefs. New entrants have mean beliefs equal to the mean of employers currently in the market. See Figure 1 for other parameter choices.

6 Appendix 1 - Proofs of Propositions 1-3

6.1 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.

6.2 Proposition 2

If $\mu \in \mathbb{R}$ is drawn from $\mu_B \sim N(\mu_0, 1/\tau_0)$ and hiring signals $x_1, ..., x_K$ are i.i.d draws from $X \sim N(\mu, 1/\tau)$, Doob (1949) shows under more general conditions that $\mu_B \rightarrow^d \mu$ as $K \rightarrow \infty$. 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 $\lambda_{jt} \rightarrow 0$ as $t \rightarrow \infty$.

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

$$\Delta V_{jt} = V(\psi'_{S_{t+1,j}}, w_{B,t+1}(\Psi_{t+1})) - V(\psi_{S_{t+1,j}}, w_{B,t+1}(\Psi_{t+1})) \text{ and}$$
$$\Delta f_{jt} = \mu - E[\mu_B | S_{tj}].$$

Employer j hires from group A only if $\beta \Delta V_{jt} - \Delta f_{jt} \leq w_{Bt}(\Psi_t) - w_A$. Moreover, market clearing requires that a subset of employers hire from group B for almost all of whom $\lambda_{jt} \rightarrow 0$ as $K \rightarrow \infty$ and $\lambda_{jt} \geq w_{Bt}(\Psi_t) - w_A$. Thus, $w_A \geq w_{Bt}(\Psi_t)$ asymptotically. Additionally, since the value of information ΔV_{jt} is weakly positive, then $\Delta f_{jt} > 0$ for this group. $\Delta f_{jt} > 0$ implies that $E[\mu_B | S_{tj}] < \mu$. 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.

6.3 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 $e_{B,t+1} \subset E_{Bt}$ with $\lambda_{j,t+1} < w_{Bt} - w_A \leq \lambda_{jt}$.[34] Suppose $w_{B,t+1} \geq w_B$, then $E_{B,t+1} \subset E_{Bt}$ and the labor market doesn't clear. Thus, $w_{B,t+1}$ 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 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

[34] This does not rely on unbounded signals. The continuum assumption ensures that a mass of employers is arbitrarily close to the cutoff such that even a slightly lower signal pushes them below.

$F_B - \varepsilon$ of employers currently hiring from Group B have value of learning smaller than ε and will hire from Group B in the limit.[35] 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 $\mu_B|\{S_{t'j} \setminus S_{tj}\}$ with the same parameters. Given that $E[\mu_B|\{S_{t'j} \setminus S_{tj}\}]$ 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[36] 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 $c < w_A$.

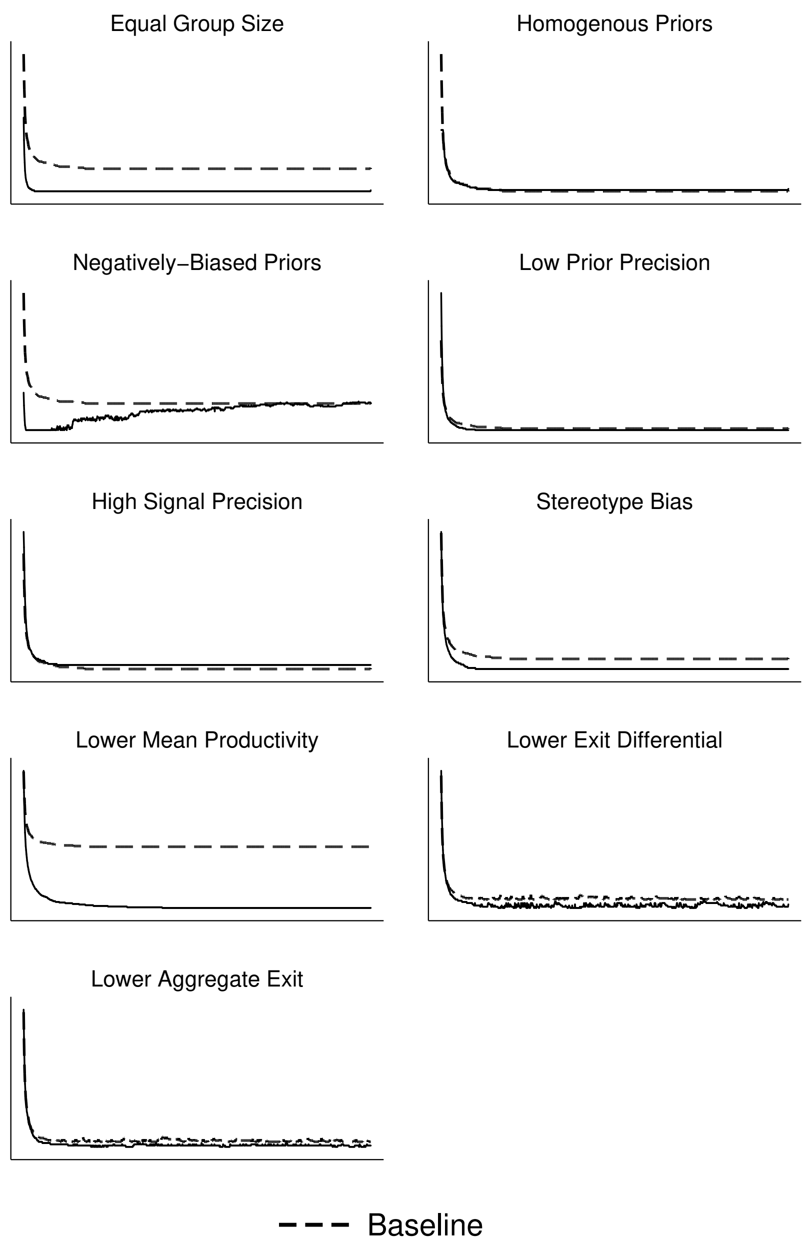
[35]This is because the value of learning and the probability that an employer currently hiring from group B falls below the cutoff next period go to 0 asymptotically.

[36]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.

7 Appendix 2 - Additional Simulations and Comparative Dynamics

The expected size of the wage gap is determined by the exogenous parameters of the model as shown in Figure A2. 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 group B is objectively less productive than employers initially believe, their long-run wage will lie below their true average productivity. 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. Assuming unbiased priors, a higher prior precision or lower variance in productivity increases the wage of group B. Assuming common 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. Lastly, with entry and exit of employers, when new employers hold unbiased priors, a lower exit rate differential for employers with negatively-biased beliefs leads to a lower wage for group B, as does a lower aggregate exit rate.

Figure A2: Wage Gap and Model Parameters



Equal Group Size refers to group B being of equal size to group A. 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 biased employers 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.

8 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_{tj} 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

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

The hiring decision hinges on the expected productivity of both groups of workers, which is decreasing in negative hiring experiences (lower draws than expected from choosing group A) and increasing otherwise. As such, hiring decisions, market clearing conditions and wage setting are unchanged, along with Proposition 1. Proposition 2 follows under regularity conditions directly applicable to $G(\cdot)$ and $h(\cdot)$ described in Section 7.2. Proposition 3 follows from assumptions made on $G(\cdot)$ as well as Propositions 1-2.