Early-Career Discrimination: Spiraling or Self-Correcting?

Arjada Bardhi[*] Yingni Guo[†] Bruno Strulovici[‡]

June 30, 2020

Abstract

Do workers from social groups with comparable productivity distributions obtain comparable lifetime earnings? We study how a small amount of early-career discrimination propagates over time when workers' productivity is revealed through employment. *Breakdown learning* environments that track on-the-job failures grant a disproportionately large advantage to marginally more favored groups, whereas *breakthrough learning* environments that track successes guarantee comparable earnings to groups of comparable productivity. This discrepancy persists with large labor markets, flexible wages, inconclusive signals, and misspecified employer beliefs. Allowing for investment in productivity exacerbates inequality between groups under breakdown learning.

JEL: D83, J71, C78

Keywords: early-career statistical discrimination, star jobs, guardian jobs, spiraling discrimination, self-correcting discrimination

1 Introduction

Young workers enter the labor market with uncertain productivity levels. To cope with this uncertainty, employers have been documented to rely on observable characteristics — such as a worker's gender or race — as statistical proxies for the worker's productivity.[1] Such early-career statistical discrimination determines who makes the first cut when opportunities are

[*]Department of Economics, Duke University. Email: arjada.bardhi@duke.edu.

[†]Department of Economics, Northwestern University. Email: yingni.guo@northwestern.edu.

[‡]Department of Economics, Northwestern University. Email: b-strulovici@northwestern.edu.

[1]Discrimination in hiring practices has been empirically documented in Goldin and Rouse (2000), Pager (2003), Bertrand and Mullainathan (2004), and other studies surveyed by Bertrand and Duflo (2017).

scarce: workers from ex-ante more productive social groups may be systematically prioritized even when the initial productivity difference with other groups is small.

Does the impact of such early-career discrimination on workers' earnings vanish or intensify over time? One conjecture is that social groups of comparable productivity obtain comparable lifetime earnings: an employer learns more about workers' productivity after observing their on-the-job performance and reallocates opportunities accordingly. An opposite conjecture is that comparable groups fare drastically differently: early opportunities to perform are pivotal in a worker's career. Without these early steps, it is hard to climb the ladder.

This paper shows that the right conjecture crucially depends on how employers learn about workers' productivity. In environments that track on-the-job successes, early-career discrimination has minor consequences for workers' lifetime earnings and employment opportunities. In environments that track on-the-job failures, by contrast, early-career discrimination significantly affects workers' prospects. Its adverse effect on discriminated workers increases in the size of *both* the favored groups and the discriminated ones. Our analysis thus suggests a classification of learning environments that predicts whether the impact of early-career discrimination vanishes or gets amplified over time.

The contrast between learning environments persists when workers can invest in their productivity and, perhaps counterintuitively, when wages are flexible. Opportunities to invest in productivity magnify the difference in the groups' productivities in learning environments that track failures, but not so in those that track successes. This leads to a tradeoff between efficiency and equality: learning environments that reduce the impact of early-career discrimination also tend to stifle the expected productivity of employed workers. When wages are flexible, a possibility that we formalize in a dynamic two-sided matching model, comparable groups continue to have drastically different earnings in environments that track failures.

Our analysis focuses on labor markets in which: (i) workers from distinct groups compete for *scarce* tasks, (ii) employers learn about a worker's productivity only if the worker performs a task, and (iii) groups have comparable productivity distributions.[2] Sarsons (2019) studies

[2]These stylized features tractably capture a more general environment in which (i') some tasks are more desirable than others and desirable tasks are in limited supply, which may be due to limited industry capacity or diseconomies of scale, (ii') workers who perform desirable tasks reveal more about their ability to perform these tasks than do workers who are either employed in other tasks or unemployed, and (iii') groups have very different productivity distributions. Our focus on comparable groups provides a particularly stark illustration of the role played by the learning environment in the dynamics of statistical discrimination.

one such market, in which male and female surgeons compete for referrals from physicians. Physicians learn about surgeons' abilities from their performed surgeries. Sarsons (2019) documents that male and female surgeons have comparable abilities: in her sample, female surgeons have a lower average ability, but the difference is small.[3] We seek to quantify the cumulative advantage resulting from such a small initial difference.

To introduce our results in their simplest form, we begin the analysis with a *baseline model* that features one employer and two workers identified by their respective social groups, a and b. We then generalize the analysis to a *dynamic two-sided matching model* with multiple workers in each group and multiple employers. In the baseline model, a worker's productivity is either high or low and worker a is ex ante more likely than worker b to have high productivity. At each instant, the employer allocates the task to one of the two workers — just like a physician choosing a surgeon for referral in the setting of Sarsons (2019) — or takes an outside option if the expected productivity of both workers is too low. The employer's flow payoff is increasing in the productivity of the employed worker. A worker benefits from being allocated the task regardless of his productivity.

The employer learns about a worker's productivity from observing the worker's performance on the task. We contrast two learning environments: *breakthrough* and *breakdown*. In the breakthrough environment, a high-productivity worker generates successes, or "breakthroughs," at randomly distributed times and a low-productivity worker generates no successes. In the breakdown environment, a low-productivity worker generates failures, or "breakdowns," at randomly distributed times, whereas a high-productivity worker does not.[4]

The learning environment may be interpreted as an intrinsic feature of the job considered. Breakthrough and breakdown environments correspond, respectively, to "star jobs" and "guardian jobs" as conceptualized by Jacobs (1981) and Baron and Kreps (1999). Scientific researchers and high-stakes salesmen are examples of star jobs; routine surgeons, airline pilots, and prison guards are examples of guardian jobs.[5]

In both learning environments, the employer first hires worker a, who is more productive

[3] See Section 2.2.2 and Figure 2 in Sarsons (2019).

[4] In a more general model, considered in Section 7.1, low-productivity workers also produce breakthroughs, but at a lower frequency. Similarly, our results are qualitatively unchanged when high-productivity workers in the breakdown environment generate breakdowns at a lower frequency than low-productivity workers.

[5] Bose and Lang (2017) argue that most nonmanagerial jobs are "guardian jobs" and derive the optimal monitoring policy for such jobs. We focus instead on the implications of guardian jobs for the propagation of small amounts of early-career discrimination and contrast them with "star jobs."

in expectation. However, subsequent task allocations differ drastically across environments. In the breakthrough environment, worker a's expected productivity declines gradually in the absence of a breakthrough, until it reaches worker b's expected productivity. From this point onward, the employer splits task allocations equally between workers until either one worker generates a breakthrough, or workers' expected productivity becomes so low that the employer prefers to take her outside option. The "grace period" over which the task is allocated exclusively to worker a reflects the difference in the employer's prior beliefs about the two workers. The smaller this belief difference, the shorter the grace period for worker a, and the smaller the first-hire advantage that worker a obtains due to his group belonging. As the belief difference shrinks to zero, so does the advantage of worker a. Hence, the breakthrough environment is *self-correcting*.

In the breakdown environment, by contrast, the absence of a breakdown from worker a makes the employer more optimistic about a's productivity. Therefore, the employer allocates the task exclusively to worker a until a breakdown arrives. Worker b is granted a chance to perform the task and demonstrate his productivity only if worker a has low productivity *and* misperforms the task. As a result, worker b's expected payoff is only a fraction of worker a's. Even if group a's productivity distribution is only slightly superior ex ante, this small initial difference turns into a large payoff advantage in the breakdown environment. Such *spiraling* effect, which does not disappear even as learning becomes very fast, can explain why societies struggle to eliminate inequality in labor markets.

The contrast between the two environments is even sharper when workers can invest in their productivity before entering the labor market. In the breakdown environment, slightly different groups invest in significantly different amounts *in every equilibrium*. Workers' payoff inequality is higher than in the baseline model, as access to investment disproportionately benefits the post-investment favored group. In the breakthrough environment, by contrast, there exists generically an equilibrium in which comparable groups invest in comparable amounts and obtain comparable payoffs. Hence, the self-correcting property of breakthroughs persists in the presence of investment opportunities.

When learning is sufficiently fast, the breakdown environment is marked by greater polarization in investment incentives. Across all equilibria, the post-investment favored worker invests strictly more under breakdowns than under breakthroughs, whereas the post-investment discriminated worker invests strictly less. When investment is sufficiently effective, *the employer prefers the breakdown environment*, since it incentivizes more investment by the post-investment favored worker. This result points to a novel tradeoff between efficiency for the employer and equality between workers.

We explore this contrast further in a large market with many workers from each group and many employers. First, we show that the key determinant of spiraling in the breakdown environment is the *relative* scarcity of tasks. Moreover, the scarcer tasks are relative to workers of either group, the greater is the inequality between groups. This implies that, while all groups suffer from a decrease in labor demand during economic downturns, discriminated groups suffer disproportionately more.

There is a widespread view that wage flexibility eliminates inequality between groups of comparable productivities. To evaluate this view, we introduce flexible wages in the large market just described. From a methodological standpoint, we develop a dynamic two-sided matching framework that incorporates learning, and show that the essentially-unique stable stage-game matching is *dynamically stable* (Ali and Liu (2020)).

Contrary to this view, wage flexibility does not fix spiraling in the breakdown environment. Flexible wages cannot eliminate the delay in employment experienced by group b. This delay further implies that the employer learns more about a-workers than b-workers. Breakdown learning thus results in a non-vanishing wage gap between groups of almost identical productivity. Figure 1 illustrates this gap: the dashed curve is the wage path for an average a-worker, and the solid curve is for an average b-worker. The wage gap persists for a substantial amount of time. Remarkably, if tasks are sufficiently scarce, the gap persists even in the very long run.

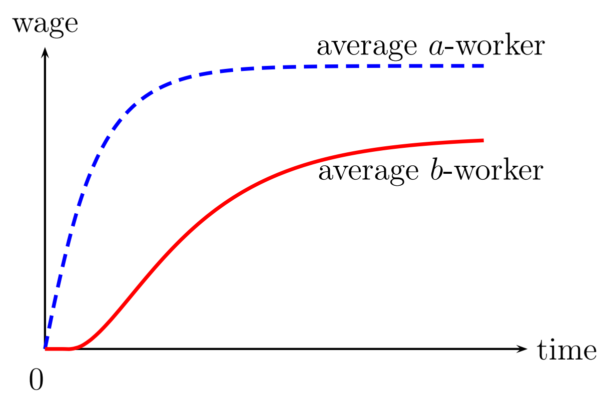

Figure 1: Wage paths under breakdowns for groups of comparable productivity

This is consistent with the persisting gender pay gap among surgeons (Lo Sasso et al. (2011), Sarsons (2019)). In line with our emphasis of early-career discrimination, a recent statement by the Association of Women Surgeons finds that "[T]he disparities women face in compensation at entry level positions lead to a persistent trend of unequal pay for equal work throughout the course of their careers."[6] We provide a learning-based mechanism that explains part of this disparity in wage paths.

The contrast between environments is still present with flexible wages: the wage paths of the two groups are arbitrarily close in the breakthrough environment. Lang and Lehmann (2012) observe that it is challenging to explain the simultaneous presence of large racial wage gaps in some sectors and the absence of such gaps in other sectors. By studying the wage gap across learning environments, our model is a step towards explaining such wide variation. We predict that, fixing everything else, breakthrough-like sectors have the tendency to exhibit much smaller wage gaps than breakdown-like ones.

Lastly, besides objective productivity differences, prejudice may be another cause of early-career discrimination. Specifically, different groups have the same expected productivity, but employers believe that one group is inferior. Such prejudice could be due to inaccurate stereotypes or due to inaccurate information about the candidate pool. The contrast between the two learning environments extends to such a setting as well, as shown in Section 7.2. In particular, under breakdown learning, prejudice among employers, even when very mild, carries dire consequences for the discriminated group.

1.1 Related literature

Our paper contributes to the literature on statistical discrimination. Fang and Moro (2011) provide a survey of theories in this literature. Phelps (1972) and the literature that followed it (e.g., Aigner and Cain (1977), Cornell and Welch (1996), and Fershtman and Pavan (2020)) assume that there is a significant, exogenous difference between social groups. Inequality between groups arises due to this difference. By contrast, Arrow (1973) and the literature that followed it (e.g., Coate and Loury (1993), Foster and Vohra (1992), Moro and Norman (2004), and Gu and Norman (2020)) assume no exogenous difference between groups. Inequality arises because groups coordinate on different equilibria or specialize in different roles

[6]For more, see AWS's Statement on Gender Salary Equity.

within an equilibrium.[7]

Our approach differs from both strands. First, we assume comparable groups, so the group difference is vanishingly small. In the setup of Phelps (1972) and subsequent models, inequality across groups disappears as the difference between groups vanishes, whereas our model highlights that a vanishingly small difference can snowball into a large payoff gap. Second, we differ from Arrow (1973) in that inequality across groups is not due to the existence of multiple equilibria. Unlike most papers in both strands, we model group interaction by letting groups compete for scarce tasks. From this standpoint our paper is related to Cornell and Welch (1996) and Moro and Norman (2004). Their task allocation occurs in one shot, whereas we explore repeated allocation.

Our results also contribute two key insights to the literature on cumulative discrimination (e.g., Blank, Dabady and Citro (2004), Blank (2005)). First, the learning environment is critical for whether cumulative discrimination arises. Second, the prospect of future cumulative discrimination casts a long shadow on workers' investment in productivity.

The paper also contributes to the literature on employer learning (e.g., Farber and Gibbons (1996), Altonji and Pierret (2001)). We think of the learning environment as an intrinsic feature of occupations or organizational ranks. On this point our work relates to Mansour (2012), Altonji (2005), Lange (2007), and Antonovics and Golan (2012). They assume that the arrival rate of signals varies across occupations, but the learning environment is otherwise fixed. Instead, we allow the nature of learning to differ across occupations, and demonstrate the importance of this variation.

Our analysis leverages the tractability of Poisson bandits, which have been used widely in strategic experimentation models (e.g., Keller, Rady and Cripps (2005), Keller and Rady (2010), Strulovici (2010), Keller and Rady (2015)).[8] In our setting, the employer is the bandit operator and the workers are the bandit arms. We also explore competing workers' incentives to invest in productivity, which corresponds to the quality of the arms being endogenously determined. In modeling bandit arms as strategic players, our paper relates to Bergemann and Valimaki (1996), Felli and Harris (1996), and Deb, Mitchell and Pai (2019).

[7]Blume (2006) and Kim and Loury (2018) extend the static setup of Coate and Loury (1993) to incorporate generations of workers. By contrast, we examine a single generation of long-lived workers whose productivities are revealed gradually only when assigned tasks.

[8]Other applications include moral hazard (e.g., Bergemann and Hege (2005), Hörner and Samuelson (2013) Halac, Kartik and Liu (2016)), team problem (e.g., Bonatti and Hörner (2011)), delegation (e.g., Guo (2016)), contest design (e.g., Halac, Kartik and Liu (2017)), etc.

In contrast to our investment analysis, they all assume that the qualities of the bandit arms are exogenously given.[9] An exception is Ghosh and Hummel (2012), in which arms' qualities are endogenous.

Section 3 identifies the key contrast between breakdown and breakthrough environments. Section 4 analyzes the role of investment in productivity. Section 5 generalizes the model to arbitrary numbers of workers and employers. It studies how the relative size of each group affects the extent of discrimination. Section 6 endogenizes the payoff that workers get from employment, and shows that allowing for flexible wages does not overturn the discrepancy between breakthrough and breakdown environments. Section 7 explores the robustness of our results with respect to more general learning environments and misspecified beliefs. Section 8 concludes.

2 Baseline model

Players and types. Time $t \in [0, \infty)$ is continuous, and the discount rate is $r > 0$. There is one employer ("she") and two workers (each "he"). Workers belong to one of two social groups, a or b. We refer to the worker from group $i \in \{a, b\}$ as worker i.

Before time $t = 0$, each worker draws his type independently from the other worker's type. Worker i's type θ_i is either high ($\theta_i = h$) or low ($\theta_i = \ell$). The probability that worker i is a high type is $p_i \in (0, 1)$. The employer knows the prior belief for each group (p_a, p_b), but she does not observe workers' types. We assume that worker a is ex-ante more productive: $p_a > p_b$. We have a focus on settings in which p_b is close to p_a.

Task allocation. At each $t \geqslant 0$, the employer allocates a task either to one of the two workers or to a safe arm. Allocating the task to the safe arm can be interpreted as the employer resorting to her outside option.

The employer obtains a flow payoff $v > 0$ if she allocates the task to a high-type worker and zero if she allocates the task to a low-type one. If she allocates the task to the safe arm, she earns a flow payoff $s \in (0, v)$. The employer's payoffs are observed at the end of the horizon.

A worker obtains a flow payoff $w > 0$, whenever he is assigned the task. Otherwise, his

[9] Another commonality between this paper and Felli and Harris (1996) is that both use a bandit setting to model labor market learning.

flow payoff is zero. This w could be interpreted as the fixed wage for a worker to perform a task. Without loss of generality, we normalize w to one. In Section 6, we extend our analysis to settings with flexible wages.

Learning by allocating. Learning about a worker's type proceeds via Poisson signals. If worker i is allocated the task over interval $[t, t+\mathrm{d}t)$ and his type is θ_i, a public signal arrives with probability $\lambda_{\theta_i}\mathrm{d}t$. With complementary probability $(1-\lambda_{\theta_i}\mathrm{d}t)$, no signal arrives.

Thus, a *learning environment* is characterized by a pair of type-dependent arrival rates $(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$. Based on whether the arrival of a signal reveals a high or a low type, we distinguish between two learning environments:

(i) *breakthrough environment*: a signal is a breakthrough if $\lambda_h > 0 = \lambda_\ell$;

(ii) *breakdown environment*: a signal is a breakdown if $\lambda_\ell > 0 = \lambda_h$.

A breakthrough perfectly reveals a high-type worker, whereas a breakdown perfectly reveals a low-type one. In Section 7.1, we extend our analysis to inconclusive breakthrough environments ($\lambda_h > \lambda_\ell > 0$) and inconclusive breakdown environments ($\lambda_\ell > \lambda_h > 0$).[10]

As discussed in the introduction, we can interpret the learning environment as an intrinsic feature of the job. We can also interpret it as an intrinsic feature of how performance is evaluated. The breakthrough environment tracks overperformance (breakthroughs) relative to expected performance (no signals), whereas the breakdown environment tracks underperformance (breakdowns).

We let $\underline{p}$ denote the belief below which the employer switches to the safe arm:

$$\underline{p} := \begin{cases} \dfrac{rs}{(r+\lambda_h)(v-s)+rs} & \text{if } \lambda_h > 0 = \lambda_\ell \\ \dfrac{(r+\lambda_\ell)s}{r(v-s)+(r+\lambda_\ell)s} & \text{if } \lambda_\ell > 0 = \lambda_h. \end{cases}$$

The threshold $\underline{p}$ is lower than the myopic threshold s/v. We assume that $p_i > \underline{p}$ for $i \in \{a, b\}$, so in either environment the employer prefers to experiment with both workers before turning

[10] In our formulation, the employer learns through signals rather than her payoffs. This is equivalent to an alternative formulation in which the employer learns through observable payoffs. In this alternative formulation, in the breakthrough environment type h generates a lump-sum benefit at arrival rate $\lambda_h > 0$ and the safe arm generates a flow payoff. In the breakdown environment, type ℓ generates a lump-sum cost at rate $\lambda_\ell > 0$ and the safe arm generates a flow cost. Our formulation makes it easier to compare the two learning environments.

to the safe arm.

3 Benchmark: A stark contrast

This section establishes a stark contrast between the two learning environments. We analyze the workers' expected lifetime payoffs given the employer's optimal hiring behavior. Our focus is on groups of comparable expected productivities. Hence, we compare the workers' payoffs as their expected productivities converge and address how this comparison depends on whether the signal takes the form of a breakthrough (Section 3.1) or a breakdown (Section 3.2).

3.1 Self-correction under breakthroughs

In the breakthrough environment, the arrival of a signal conclusively proves that the worker is a high type. In the absence of a breakthrough the employer becomes more pessimistic about a worker's type. Proposition 3.1 establishes a *self-correcting property* of breakthrough learning: a small difference in prior beliefs can only result in a small payoff advantage for worker a.

The employer's optimal strategy consists in allocating the task to the worker with the higher belief at each point in time. Since $p_a > p_b$, the employer allocates the task to worker a first. Because the belief about worker a's type drifts down for as long as no breakthrough arrives, worker a is effectively given a grace period $[0, t^*)$ to perform. Here, t^* measures how long it takes for the belief about worker a's type to drift down from p_a to p_b in the absence of a breakthrough. If worker a generates a breakthrough before t^*, the employer allocates the task only to him thereafter. Otherwise, starting from t^*, the employer splits the task equally among workers until the belief drops down to $\underline{p}$, so the workers obtain the same continuation payoff starting from t^*. The hiring dynamics therefore go through two distinct phases: a first phase during which worker a is hired exclusively, and a second phase during which the two workers are treated symmetrically.

Importantly, as p_b gets close to p_a, the grace period $[0, t^*)$ shrinks to zero. The probability that worker a generates a breakthrough before t^* converges to zero as well. Hence, the two players obtain similar expected payoffs.

Proposition 3.1 (Self-correcting property of breakthroughs)**.** *As $p_b \uparrow p_a$, the expected payoff of worker b converges to that of worker a.*

Proof. The employer allocates the task first to worker a for $t \in [0, t^*)$, where t^* is the time that it takes for p_a to drop to p_b:

$$\frac{p_a e^{-\lambda_h t^*}}{p_a e^{-\lambda_h t^*} + 1 - p_a} = p_b \quad \Longrightarrow \quad t^* = \frac{1}{\lambda_h} \log \frac{p_a(1-p_b)}{(1-p_a)p_b}. \tag{1}$$

After that, the employer splits the task equally between the two workers until either (i) a breakthrough arrives, or (ii) the belief about the workers' types hits $\underline{p}$. If a breakthrough arrives, the employer allocates the task only to the worker who generated the breakthrough thereafter.

We let $U_i(p_a, p_b)$ be worker i's payoff given beliefs (p_a, p_b). Note that $U_a(p,p) = U_b(p,p)$ for any $p \in (\underline{p}, 1)$. Over interval $[0, t^*)$, worker a generates a breakthrough with probability $p_a \left(1 - e^{-\lambda_h t^*}\right)$. If a breakthrough arrives, worker a's payoff is 1. If it does not arrive, worker a's payoff consists of $1 - e^{-rt^*}$, the flow payoff from $[0, t^*)$, and $U_a(p_b, p_b)$, the continuation payoff from time t^* onward. Worker a's total payoff is

$$p_a \left(1 - e^{-\lambda_h t^*}\right) + \left(1 - p_a + p_a e^{-\lambda_h t^*}\right) \left(1 - e^{-rt^*} + e^{-rt^*} U_a(p_b, p_b)\right).$$

Worker b gets continuation payoff $U_b(p_b, p_b)$ at time t^* if and only if no breakthrough occurs over $[0, t^*)$:

$$\left(1 - p_a + p_a e^{-\lambda_h t^*}\right) e^{-rt^*} U_b(p_b, p_b).$$

As $p_b \uparrow p_a$, $t^* \to 0$. Therefore, the two workers' payoffs are equal in the limit. ■

3.2 Spiraling under breakdowns

We now turn to the breakdown environment, in which a signal conclusively proves that the worker is a low type. As long as a worker generates no breakdown, the employer becomes more optimistic that his type is high. She first allocates the task to worker a. In the absence of a breakdown, the employer continues hiring him. If a breakdown is realized, the employer turns to worker b immediately. If worker b generates a breakdown as well, she resorts to the safe arm thereafter.

Proposition 3.2 establishes a *spiraling property* of breakdown learning: even if p_b is just shy of p_a, worker a has a substantially higher payoff than worker b does. In fact, worker a obtains the same payoff as if worker b did not exist: he is the first to be hired and remains so until he generates a breakdown. This stands in contrast to the breakthrough environment, where worker a loses his preferential status if he fails to generate a breakthrough within a given time window.

Proposition 3.2 (Spiraling property of breakdowns)**.** *As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a approaches*

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r} < 1.$$

Proof. If worker a is a high type, his payoff is 1. If he is a low type and the first breakdown arrives at t, his payoff is $(1-e^{-rt})$. The arrival time t follows density $\lambda_\ell e^{-\lambda_\ell t}$. Hence, worker a's expected payoff is:

$$p_a + (1-p_a)\frac{r}{\lambda_\ell + r}. \tag{2}$$

Worker b's payoff if the employer starts hiring him at time t is:

$$e^{-rt}\left(p_b + (1-p_b)\frac{r}{\lambda_\ell + r}\right).$$

Conditional on worker a being a low type, this time t is distributed according to density $\lambda_\ell e^{-\lambda_\ell t}$. Hence, worker b's expected payoff is:

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(p_b + (1-p_b)\frac{r}{\lambda_\ell + r}\right). \tag{3}$$

■

At the heart of spiraling is the fact that the delay that worker b faces does not depend on how close p_b is to p_a. The payoff ratio in Proposition 3.2 has two components: (i) $(1-p_a)$ reflects the fact that worker b obtains a chance only if worker a is a low type, and (ii) $\lambda_\ell/(\lambda_\ell + r)$ reflects the expected time it takes for a's low type to be revealed. Moreover, even as learning becomes instantaneous — i.e., as $\lambda_\ell \to +\infty$ — this payoff ratio approaches $(1-p_a)$ rather than one. Being the second hire is detrimental to worker b even when the employer learns very fast: worker b never obtains a chance if worker a is a high type.

Since groups have similar productivity distributions, even if the employer were blind to workers' group belonging and treated the workers equally, her payoff would not be much lower than when she observes group belonging. In the limit as $p_b \uparrow p_a$, her payoff would be the same in both cases. Therefore, making it more difficult for the employer to observe group belonging would equalize workers' payoffs without making the employer worse off.[11]

4 Investment in productivity

This section allows workers to invest in their types before they enter the labor market, and explores the implications of this investment opportunity. Such implications are a priori unclear: access to investment might level the playing field or it might amplify the productivity gap between workers. Yet again, it turns out that the learning environment plays a key role.

Section 4.2 establishes that access to investment does not disturb the self-correcting property of the breakthrough environment. In the breakdown environment, however, investment exacerbates spiraling, in the sense that it makes the workers' payoffs more unequal. Furthermore, Section 4.3 shows that when learning is sufficiently fast, breakdown learning leads to more polarized investment behavior across workers than breakthrough learning does. Hence, the post-investment productivity gap is larger under breakdowns than under breakthroughs.

Formally, the investment stage occurs before time $t = 0$. It takes three steps: (i) workers draw their pre-investment types independently according to probabilities (p_a, p_b); (ii) a low-type worker draws his cost of investment $c \in [0, 1]$ according to the cumulative distribution function F and decides whether to invest; (iii) if he invests, he pays cost c and his type improves to high with probability $\pi \in (0, 1)$. Each worker's investment cost, investment decision, and post-investment type are observed by himself only. We assume that F is continuously differentiable and strictly increasing.

Workers enter the labor market at time $t = 0$ and encounter the employer. Subsequently at each $t \geqslant 0$, the employer allocates a task either to one of the two workers or to a safe arm.

[11] In a study of blind hiring practices, Goldin and Rouse (2000) show that blind orchestra auditions substantially increased the likelihood that female musicians advanced to the final round.

4.1 Preliminary analysis

Before turning to the results, we first characterize the set of equilibria in the investment game. Let (q_a, q_b) denote the employer's belief about each worker after the investment stage. The employer follows an optimal hiring strategy given this belief pair. We let $B_i(q_a, q_b)$ denote the benefit of investment for a low-type worker i:

$$B_i(q_a, q_b) := \pi \left(U_i(h; q_a, q_b) - U_i(\ell; q_a, q_b) \right),$$

where $U_i(\theta_i; q_a, q_b)$ is the expected payoff of worker i with post-investment type θ_i given employer's optimal hiring strategy for (q_a, q_b). A low-type worker invests if and only if the benefit of doing so exceeds his realized cost, drawn according to F. Hence, in any equilibrium a worker's investment strategy takes a threshold form. We let c_i be the cost threshold below which worker i invests.

An equilibrium is characterized by a pair of cost thresholds (c_a, c_b) and a pair of post-investment beliefs (q_a, q_b) such that:

1. the employer chooses an optimal hiring strategy given (q_a, q_b);
2. (c_a, c_b) are the workers' best responses to the hiring strategy induced by (q_a, q_b), i.e., $c_i = B_i(q_a, q_b)$;
3. (q_a, q_b) is consistent with workers' investment strategy (c_a, c_b), i.e., $q_i = p_i + (1 - p_i)F(c_i)\pi$.

A feature common to both learning environments is that if the employer believes that worker i is more productive than worker $-i$ after the investment stage, then i's benefit from investment is strictly higher than that of $-i$.[12] This is because worker i is the first to be allocated the task: by investing, he might avoid a breakdown in the near future or increase the chance of a breakthrough within the grace period allotted exclusively to him.

Lemma 4.1 (Post-investment favored worker has higher benefit of investment)**.** *In both learning environments, if $q_i > q_{-i}$ then $B_i(q_a, q_b) > B_{-i}(q_a, q_b)$. For each i, $B_i(q_a, q_b)$ is*

[12] In the breakthrough environment, if $q_a = q_b = q$ the employer optimally splits his task between the workers, hence $B_a(q, q) = B_b(q, q)$. This is also the case in the breakdown environment, assuming that the employer randomizes equally between workers at $t = 0$ if $q_a = q_b$.

continuously differentiable in the breakthrough environment, but it is not continuous at $q_a = q_b$ in the breakdown environment.

Proof of Lemma 4.1. Suppose $q_a > q_b$, and let $\mu_\ell := \lambda_\ell / r$. Worker i's benefit from investment is $B_i(q_a, q_b) = \pi\Big(U_i(h; q_a, q_b) - U_i(\ell; q_a, q_b)\Big)$. We first show the inequality for the breakdown environment. Then:

$$U_a(\theta_a; q_a, q_b) = \begin{cases} 1 & \text{if } \theta_a = h \\ \dfrac{1}{\mu_\ell + 1} & \text{if } \theta_a = \ell \end{cases}, \qquad U_b(\theta_b; q_a, q_b) = \begin{cases} \dfrac{\mu_\ell(1 - q_a)}{\mu_\ell + 1} & \text{if } \theta_b = h \\ \dfrac{\mu_\ell(1 - q_a)}{(\mu_\ell + 1)^2} & \text{if } \theta_b = \ell. \end{cases}$$

It follows that if $q_a > q_b$ then

$$B_a(q_a, q_b) = \pi \frac{\mu_\ell}{\mu_\ell + 1} > B_b(q_a, q_b) = \pi \left(\frac{\mu_\ell}{1 + \mu_\ell} \right)^2 (1 - q_a).$$

Hence the benefit to the worker with the higher post-investment belief is strictly higher. Again, the benefit of investment for worker i is:

$$B_i(q_a, q_b) = \begin{cases} \pi \dfrac{\mu_\ell}{\mu_\ell + 1} & \text{if } q_i > q_{-i} \\ \pi \left(\dfrac{\mu_\ell}{1 + \mu_\ell} \right)^2 (1 - q_{-i}) & \text{if } q_i < q_{-i}. \end{cases}$$

Hence, the benefit of investment for worker i is discontinuous at $q_i = q_{-i}$.

The proof the breakthrough environment follows similar steps but is more algebraically involved. It is contained in Appendix A. ■

The post-investment favored worker i has stronger incentives to invest. This, in turn, rationalizes the employer's ranking of worker i over worker $-i$ in equilibrium. Not surprisingly, such a self-fulfilling force — which is also present in Coate and Loury (1993) — leads to multiple equilibria. In fact, when the prior gap is sufficiently small, there exist equilibria in which worker b overtakes worker a and thus becomes post-investment favored. Therefore, investment can reverse the initial ranking of groups. Faced with such multiplicity of equilibria, we focus on the comparison of the equilibrium *sets* across the two learning environments.

Lastly, when $p_a \neq p_b$, there cannot exist any equilibrium for which $q_a = q_b$. If such an equilibrium existed, investment would provide the same benefit of investment to both workers. Workers would therefore use the same investment strategy. However, since their probabilities of being a high type are different (i.e., $p_b < p_a$), worker b would need to invest more in order to reach $q_a = q_b$, which contradicts workers' identical investment strategies.

4.2 Exacerbated spiraling under breakdowns

Our next proposition formalizes the notion that the self-correcting property of the breakthrough environment continues to hold. There always exists an equilibrium in which q_b converges to q_a as $p_b \uparrow p_a$. The workers' benefits from investment get arbitrarily close, and so do their investment thresholds. Hence, the payoff gap between comparable workers becomes vanishingly small. This equilibrium could either preserve the prior ranking of the workers — with worker a being post-investment favored — or reverse it.

Proposition 4.1 (Self-correction under breakthroughs)**.** *Suppose that F is weakly convex. For a generic set of parameters, as $p_b \uparrow p_a$, there exists an equilibrium in which the two workers' expected payoffs as well as their post-investment beliefs converge.*

By a "generic set of parameters", we mean that once we fix the cost distribution F and all other parameters except for (p_a, π), the set of values of $(p_a, \pi) \in \big(\underline{p}, 1\big) \times (0, 1)$ for which the proposition does not hold has measure zero.[13] The proof builds on two observations. First, when the two workers have the same prior belief (i.e., $p_a = p_b$), there always exists a symmetric equilibrium in which the workers follow the same cost threshold and thus have the same post-investment belief. Second, under breakthrough learning the benefit from investment is continuously differentiable in (q_a, q_b). We invoke the implicit function theorem to assert that when p_b is within a small neighborhood of p_a, there exists an equilibrium in which cost thresholds (c_a, c_b) and post-investment beliefs (q_a, q_b) are within a small neighborhood of those in the symmetric equilibrium.

By contrast, access to investment not only fails to tame the propensity of breakdown learning to magnify small prior differences, but it makes it worse. Across *all* equilibria,

[13] When F is non-convex, a version of the result continues to hold according to a different, more involved, notion of genericity.

the payoffs of ex ante comparable workers are even further apart than in the no-investment benchmark.

Proposition 4.2 (Exacerbated spiraling under breakdowns). *As $p_b \uparrow p_a$, in any equilibrium (q_i, q_{-i}) such that $q_i > q_{-i}$, the ratio of the expected payoff of worker $-i$ to that of worker i is at most*

$$(1-q_i)\frac{\lambda_\ell}{\lambda_\ell + r} < 1,$$

which is strictly lower than the payoff ratio in the no-investment benchmark, given by $(1-p_a)\lambda_\ell/(\lambda_\ell + r)$.

Unlike in the breakthrough environment, the benefit from investment in the breakdown environment is discontinuous in (q_a, q_b), as shown in Lemma 4.1. This difference explains why a proof similar to that in Proposition 4.1 does not work here. Nonetheless, the benefit function takes a simple form — as already previewed in the proof of Lemma 4.1 — hence characterizing the set of equilibria is straightforward. As $p_b \uparrow p_a$ there exist only two equilibria: worker a is post-investment favored in one equilibrium and worker b is post-investment favored in the other. In the limit $p_b \uparrow p_a$, these two equilibria look the same modulo the workers' identities.

As we saw in Proposition 3.2, the payoff ratio across workers in the no-investment benchmark is pinned down by p_a, the belief of worker a without investment. Because the post-investment favored worker — whoever that might be — has a strong enough incentive to invest, his post-investment belief is strictly higher than p_a. We show that for any realized investment cost, the ratio between the payoff of the post-investment discriminated worker to that of the post-investment favored worker — after factoring in the investment cost — is lower than that in the no-investment benchmark. Hence, inequality between workers increases due to the investment opportunity.

4.3 Polarization in investment behavior under breakdowns

Our discussion so far focused on the comparison of learning environments in terms of workers' payoffs. We now turn to differences in investment behavior. In a nutshell, when learning is sufficiently fast, the post-investment favored worker invests strictly more under breakdowns than under breakthroughs. The post-investment discriminated worker, by contrast, invests strictly less under breakdowns. Therefore, the breakdown environment is marked by greater polarization in workers' investment behavior.

Proposition 4.3 (Investment polarization under breakdowns)**.**

(i) Fixing all but $(\lambda_h, \lambda_\ell)$*, there exists* $\bar{\lambda} > 0$ *such that for any* $(\lambda_h, \lambda_\ell) \in \left[\bar{\lambda}, \infty\right)^2$ *and in any equilibrium, the post-investment favored worker invests strictly more in the breakdown environment than in the breakthrough one.*

(ii) Fixing all but $(\lambda_h, \lambda_\ell, \underline{p})$*, there exists* $\tilde{\lambda} > 0$ *and* $\tilde{\underline{p}} > 0$ *such that for any* $(\lambda_h, \lambda_\ell) \in \left[\tilde{\lambda}, \infty\right)^2$ *and any* $\underline{p} \leqslant \tilde{\underline{p}}$*, the post-investment discriminated worker invests strictly less in the breakdown environment than in the breakthrough one.*

When learning is sufficiently fast, investment incentives are seemingly similar across the two environments. For the post-investment favored worker, he is the first to be allocated the task and information about his type arrives quickly, so his investment incentives are quite strong. For the post-investment discriminated worker, investment incentives are less clear. On the one hand, a stronger favored worker depresses the discriminated worker's incentives to invest. On the other, due to fast learning, the discriminated worker might get a chance earlier. Despite these seeming similarities across environments, Proposition 4.3 identifies key differences between the two environments. For the result, the arrival rates need not be equal across environments: all that the result requires are sufficiently high arrival rates. Moreover, the result holds across all equilibria.

Under breakdown learning, the return from being a high type is very close to one for the post-investment favored worker. A high type gets a payoff of one, and a low type is revealed and fired almost immediately. Under breakthrough learning, by contrast, the return from being a high type is not as high. As learning becomes very fast, the grace period granted to the favored worker becomes very small. A high type is revealed with probability strictly less than one during this small grace period. This uncertainty about whether a high type will be revealed reduces the return from investment. Therefore, the post-investment favored worker is more motivated to invest in the breakdown environment.

The discriminated worker, on the other hand, has a weaker incentive to invest under breakdowns than under breakthroughs when the safe arm is sufficiently unappealing to the employer. This is due to two forces that reinforce each other. First, because of spiraling the breakdown environment already disfavors the second worker to be hired much more than the breakthrough environment does. Second, under breakdowns the discriminated worker faces an opponent that invests strictly more — as explained above — which further lowers the chance that the discriminated worker will get a shot from the employer.

An important implication of Proposition 4.3 is that when learning is sufficiently fast and π sufficiently close to 1, the employer strictly prefers the breakdown environment to the breakthrough one. That is, if she had the choice between learning environments, she would opt for the environment that magnifies small differences. Therefore, there is a tradeoff between efficiency for the employer and equality across workers. The employer prefers the breakdown environment because this environment encourages almost sure investment by the favored worker, hence she hires a high type almost surely. In the breakthrough environment, by contrast, the favored worker is a low type with positive probability after investment. Hence, the employer's payoff is bounded away from v.

Corollary 4.2. *There exists $\bar{\lambda} > 0$ and $\bar{\pi} \in (0, 1)$ such that for any $\lambda_h, \lambda_\ell \geqslant \bar{\lambda}$ and $\pi > \bar{\pi}$, the employer's payoff is strictly higher under breakdowns than under breakthroughs.*

Indeterminate ranking under slow learning. When learning is slow, on the other hand, the environments are more similar and hence the ranking can go either way. For a simple illustration, suppose $\lambda_h \to 0$ and $\lambda_\ell = \epsilon > 0$ small. Investment under breakthroughs is very close to zero for both workers, whereas investment in the breakdown environment is small but strictly positive for both. Therefore, the breakdown environment provides greater incentives to invest for such λ_h and λ_ℓ. The reverse ranking holds if $\lambda_\ell \to 0$ and $\lambda_h = \epsilon > 0$. Therefore, it is harder to attain a clear-cut comparison when the environments are qualitatively more similar.

5 Many employers and workers: The role of task scarcity

Our baseline model focused on a single employer choosing between two workers from distinct groups. We now consider a market with many employers and many workers from each group, still modeling only two groups for simplicity.[14] Comparable groups continue to have

[14]The results in this section can be extended to more than two groups.

comparable payoffs under breakthroughs, but markedly different payoffs under breakdowns. Moreover, the scarcer tasks are relative to workers, the greater is the inequality between groups in the breakdown environment.

We consider a two-sided market with a unit mass of employers, a mass of size α of a-workers, and a mass of size β of b-workers. As before, each employer has a task to allocate at each moment. An i-worker's type is high with probability p_i and is drawn independently from other workers' types. As in our baseline model, we focus on the limit $p_b \uparrow p_a$.

Moreover, we assume that $\alpha+\beta > 1$, so tasks are scarce and some workers are not initially hired. Such task scarcity is both necessary and sufficient for the payoff gap to emerge in the breakdown environment. For ease of exposition, we assume $\alpha > 1$ for the rest of this discussion. The formal analysis is relegated to Appendix B.

Broad allocation under breakthroughs. Mirroring the analysis in Section 3.1, the dynamic allocation of tasks goes through two phases. In the first phase, all a-workers take turns to perform tasks. If a worker generates a breakthrough, he "secures his job" with his current employer: the employer allocates future tasks only to this worker thereafter. For those a-workers without a breakthrough, the employers' belief drops gradually until it reaches p_b. At that point, a-workers without breakthroughs are believed to be as productive as b-workers. The allocation now enters a second phase in which remaining employers let remaining a-workers and all b-workers take turns to perform tasks. Again, those who generate breakthroughs secure their jobs with their current employers.

Breakthrough learning, hence, prompts employers to try a broad set of workers. A similar observation was made in passing by Baron and Kreps (1999) on recruitment for star jobs:

> For a star job, the costs of a hiring error are small relative to the upside potential from finding an exceptional individual. Therefore, the organization will wish to sample widely among many employees, looking for the one pearl among the pebbles.

Our focus is on the implication of such broad allocation for group inequality. Employers quickly extend their search to group b, so a b-worker's payoff converges to an a-worker's payoff as $p_b \uparrow p_a$. Therefore the self-correcting property extends to larger labor markets.

Focused allocation under breakdowns. We next show that breakdown learning leads to sluggishness in trying new workers: if a worker is hired, he remains so until he generates a breakdown. This sluggishness hurts group b disproportionally no matter how close p_b is to p_a, thus generalizing the intuition gained from Section 3.2 to larger labor markets.

At the start, a unit mass of a-workers are hired by the unit mass of employers. These workers continue to be hired as long as they do not generate breakdowns. When one of these a-workers generates a breakdown, he is replaced by a new a-worker for as long as one is available. So, b-workers wait for their turn until all of a-workers have been tried *and* sufficiently many a-workers have generated breakdowns. Crucially, this delay does not shrink as $p_b \uparrow p_a$. Therefore, the expected payoff of a b-worker remains bounded away from that of an a-worker.

Larger groups, higher inequality under breakdowns. We measure group inequality by the ratio of a b-worker's expected payoff to that of an a-worker as $p_b \uparrow p_a$. In the breakdown environment, inequality between groups increases as the size of either group increases.

First, increasing β while keeping α fixed intensifies competition within group b but does not affect the payoff of a-workers. Second, increasing α while keeping β fixed hurts both groups: it intensifies competition within group a while also increasing delay for group b. We show that increasing α hurts group b more than it hurts group a, since adding one more a-worker uniformly delays every b-worker's employment. Therefore, the larger is the labor supply from either group — i.e., the more workers there are from either group relative to the unit mass of tasks — the greater is the inequality between groups.

Proposition 5.1 (Inequality increases in task scarcity under breakdowns)**.** *Let $\alpha > 1$ and $\beta > 0$. As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker decreases in both α and β.*

This result predicts that when the scarcity of job opportunities intensifies — e.g., when labor demand falls during an economic downturn — inequality deepens. This is consistent with the observation that while all groups suffer during an economic downturn, some suffer disproportionately more.[15]

[15]Estimates from Pew Research Center show that the white-to-black and white-to-Hispanic wealth ratios were much higher at the peak of the recession in 2009 than they had been since 1984, the first year for which the U.S. Census Bureau published wealth estimates by race and ethnicity based on the Survey of Income and Program Participation.

6 Flexible wages

Thus far we assumed that a worker's wage from performing a task is fixed. A natural question is whether flexible wages can restore payoff equality in the breakdown environment. This section shows that when workers are protected by limited liability, so that wages are non-negative, both the self-correcting property of breakthroughs and the spiraling property of breakdowns still hold. In particular, wage flexibility is insufficient to prevent spiraling.

We introduce flexible wages to the large market of Section 5. Now a stage-game outcome describes not only how workers are matched to employers but also a wage for each matched pair. We call this a *stage-game matching.* Using the solution concept of Shapley and Shubik (1971), we characterize the set of stable stage-game matchings. The dynamic counterpart of a stage-game matching — a *dynamic matching* — specifies, after each history, how workers are matched to employers along with a wage for each matched pair. Adopting the solution concept of Ali and Liu (2020), we show that prescribing the stable stage-game matching after each history is dynamically stable: no worker-employer pair and no single player has a profitable one-shot deviation after any history. The formal analysis is relegated to Appendix C.

In both learning environments, employers are matched to the most productive workers after each history. In particular, there is a history-dependent marginal belief p^M: all workers with belief greater than p^M are matched and all those with belief below p^M are not. Wage takes a strikingly simple form: a matched worker of belief p is paid $(p - p^M)v$, which is the additional value that he creates relative to the marginal-belief worker. An unmatched worker is paid zero wage. As a result, all employers get the same flow profit of $p^M v$.

Why is prescribing the stable stage-game matching after each history dynamically stable? First, no employer finds it profitable to reject being matched, since his flow profit from being matched is at least as high as his payoff from a safe arm. Second, no employer-worker pair has a profitable one-shot deviation, since all employers make the same flow profit. Lastly, no worker finds it profitable to reject being matched and delay learning. The wage, $\max\{0, (p-p^M)v\}$, is convex in the employer's belief p, as shown by the green curve in Figure 2. So learning — and hence splitting the current belief into posterior beliefs — benefits the worker.

Flexible wages do not fix spiraling. A widespread view is that with flexible wages, workers of comparable beliefs obtain comparable payoffs. This would be the case if the matching occurred only once: given wage continuity, a b-worker's payoff in a one-shot matching would converge to an a-worker's payoff as $p_b \uparrow p_a$. But this intuition no longer holds when matching occurs multiple times.

To illustrate, we consider a two-period example in Figure 2. In the first period, since $p_a > p^M > p_b$, the a-worker is hired and the b-worker is not. However, since p_a is close to p_b, the a-worker's wage w_1 is very close to zero. This shows that, within a single period, workers of comparable beliefs obtain comparable payoffs.

The a-worker's performance in the first period reveals information about his type, splitting the prior belief p_a into two posterior beliefs. We let $\overline{p}_a, \underline{p}_a$ denote the posterior beliefs, so the expected wage for the a-worker in the second period is w_2. The more information is gathered about the a-worker, the further $\overline{p}_a$ and $\underline{p}_a$ are from p_a, so the higher w_2 is. Hence, due to learning about the a-worker's type, the payoff gap between workers widens in the second period.[16]

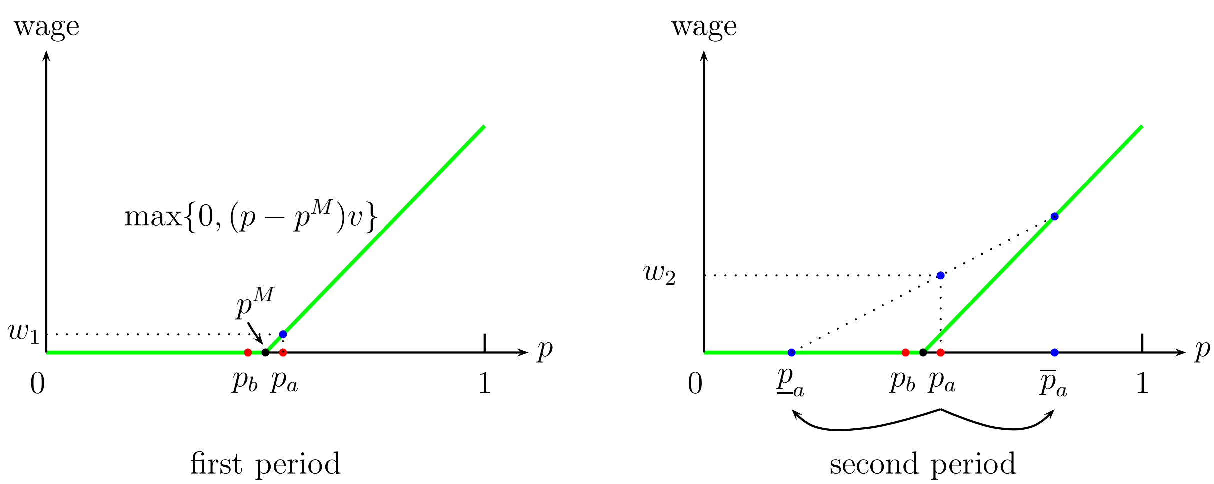

Figure 2: A two-period example

This two-period example illustrates why spiraling arises under breakdowns even with flexible wages. The group delay experienced by b-workers does not vanish even as p_b gets arbitrarily close to p_a. By the time that employers start hiring b-workers, they have already

[16]Figure 2 ignores the fact that the marginal belief p^M evolves over time. The intuition extends readily to the case with history-dependent p^M.

learned a lot about a-workers' types. Hence, the expected flow payoff of an a-worker is significantly higher than that for a b-worker. By contrast, under breakthroughs, the group delay experienced by b-workers vanishes as p_b converges to p_a, so employers have not learned much more about a-workers than about b-workers. Hence, their payoffs converge.

Persistent wage gap under breakdowns. We also obtain closed-form expressions for the expected wage path of each group as $p_b \uparrow p_a$ in the breakdown environment. In line with the persistence of spiraling, the average a-worker and the average b-worker face very different wage paths (see Proposition C.3 and Figure 3). Due to the non-vanishing delay faced by group b, the expected wage gap first expands monotonically and then it shrinks. Whether the wage gap approaches zero in the long run depends on how scarce tasks are. If task scarcity is sufficiently severe — in the sense that there are in expectation more tasks than high-type workers, so $p_a(\alpha+\beta) > 1$ — this wage gap remains bounded away from zero even as $t \to \infty$.

The wage gap at any instant is due to the combination of two factors. First, at any instant, an average a-worker has a higher chance of being hired than an average b-worker. Second, an average b-worker is hired at a later date than an average a-worker, so more is learned about this a-worker's type. Hence, conditional on being hired, the a-worker has a higher expected wage than the b-worker does.

7 Other modeling variations

7.1 Inconclusive learning environments

In our baseline model, signals are conclusive: only one type can generate the signal. More generally, a learning environment is characterized by a pair of arrival rates $(\lambda_h, \lambda_\ell)$ such that $\lambda_\theta > 0$ for both $\theta \in \{h, \ell\}$. That is, both high and low types generate signals. The environment is an *inconclusive breakthrough* one if the signal suggests a high type (i.e., $\lambda_h > \lambda_\ell$) and an *inconclusive breakdown* one otherwise (i.e., $\lambda_h < \lambda_\ell$). If $\lambda_h = \lambda_\ell$, signals are uninformative.

The self-correcting property extends to inconclusive breakthroughs, as established in Proposition D.1. Even though the employer does not assign the task to worker a indefinitely upon the realization of the first breakthrough, there is still a time window $[0, t^*)$ over which worker a should generate a first breakthrough in order to continue being allocated the task exclusively. If no breakthrough arrives during this time window, the belief about worker a's type drops to p_b, at which point both workers receive equal continuation payoffs. It continues to be the case that as $p_b \uparrow p_a$, t^* shrinks to zero and hence the probability that worker a generates a breakthrough within the time window vanishes as well. The two workers' limit payoffs are therefore equal.

The spiraling property generalizes to inconclusive breakdowns as well, provided that players are sufficiently impatient. The departure from conclusive breakdowns brings the complication that the employer might revisit workers who have generated breakdowns in the past. But as long as $p_a > p_b$, worker a is the first to be hired and stays employed in the absence of a breakdown. The expected time until the first breakdown is significant. If players are sufficiently impatient, this already leads to a significant payoff advantage for worker a.[17] Proposition D.2 presents the details.

7.2 Misspecified prior belief

Suppose that the two workers have the same probability p_{true} of being a high type, but the employer believes that worker b has a lower probability $p_{\text{mis}} < p_{\text{true}}$.[18] The spiraling property of the breakdown environment continues to hold, in the sense that even a very slight misspecification grants a large payoff disadvantage to worker b. Worker a is hired first based on the employer's misspecified belief and the workers' payoffs are still given by expressions (2) and (3) (with p_a and p_b being replaced by p_{true}).

The self-correcting property of the breakthrough environment continues to hold as well, in the sense that a slight misspecification will not have large payoff consequences for workers. Duration t^* — analogous to the grace period in (1) — now measures how long it takes for the belief about worker a's type to drift down from p_{true} to p_{mis}. As the amount of misspecification vanishes to zero, so does t^* as well. At time t^*, the true probability that worker a is a high type is p_{mis}, whereas the true probability that worker b is a high type is

[17] The sufficient condition for spiraling can be also stated in terms of arrival rates $(\lambda_h, \lambda_\ell)$ rather than the discount rate r: breakdowns need to be sufficiently infrequent, i.e., λ_h, λ_ℓ sufficiently small.

[18] Bohren et al. (2019) refer to this as "inaccurate statistical discrimination." Bohren, Imas and Rosenberg (2019) identify discrimination driven by misspecified beliefs in an experimental setting.

p_{true}. The employer, though, believes that both probabilities are p_{mis}, so she splits the task equally between workers from t^* onwards.

We let $\hat{U}_a(p_{\text{mis}}, p_{\text{true}})$ and $\hat{U}_b(p_{\text{mis}}, p_{\text{true}})$ be the continuation payoffs of worker a and b at time t^*. Because each worker gets half a task but worker a has a lower true probability of being a high type, his payoff $\hat{U}_a(p_{\text{mis}}, p_{\text{true}})$ is lower than $\hat{U}_b(p_{\text{mis}}, p_{\text{true}})$. Crucially, as p_{mis} converges to p_{true}, the two payoffs get arbitrarily close. To extend the proof of Proposition 3.1 to the misspecified-prior case, we only need to replace t^* with the new definition and replace $U_i(p_b, p_b)$ with $\hat{U}_i(p_{\text{mis}}, p_{\text{true}})$ for workers' payoffs.

Belief misspecification is quite relevant in discussions of labor-market discrimination. Lang and Lehmann (2012) provide evidence that suggests the presence of widespread mild prejudice among employers. Our results show that prejudice, even when infinitesimally mild, has very different implications in different learning environments. The breakthrough environment works well against prejudice, whereas the breakdown environment propagates it further.

8 Concluding remarks

This paper studied the payoff consequences of different learning environments for social groups of comparable productivity. We argued that whether the learning environment is closer to a breakdown environment or a breakthrough one has important implications for whether discrimination persists in the long run. Lange (2007) observed that "how economically relevant statistical discrimination is depends on how fast employers learn about workers' productive types." Our analysis adds an important missing piece to this view: what matters for statistical discrimination is not only the speed of employer learning, but also the nature of that learning.

The contrast between breakdown and breakthrough learning environments is reminiscent of the job typology introduced by Jacobs (1981) and Baron and Kreps (1999). More empirical work is needed to investigate the persistence and magnitude of discrimination in guardian jobs compared to star jobs, both in terms of the employment and promotion gaps and in terms of the wage gaps that our theory predicts. Our theoretical framework and its testable implications can guide such empirical work.

Besides the testable predictions discussed so far, one natural empirical question for which our framework can be useful is the long-lasting effects of temporary affirmative action for discriminated groups (Miller and Segal (2012), Kurtulus (2016), Miller (2017)). The empirical evidence on this question is mixed. A natural corollary of our analysis is that, in breakdown environments, if the employer is obligated to give a chance to group b early on, this dramatically improves the prospects of b-workers as they will continue to be hired even after this temporary obligation is lifted. This will not be the case in breakthrough environments.

Finally, our framework can be used to study questions that fall beyond the scope of the current paper. First, an employer may have to allocate multiple tasks that generate different learning dynamics. For instance, if an employer has both a breakthrough task and a breakdown task, how will she allocate the tasks between workers from comparable social groups? Second, in certain contexts the learning environment is an endogenous choice of the employer rather than exogenously fixed. Corollary 4.2 outlined circumstances under which the employer prefers breakdown learning. More generally, is the endogenous choice of the learning environment more likely to lead to breakdown or breakthrough learning? Third, our framework can prove useful to understand incentives for occupational segregation: workers from discriminated groups have incentives to sort into breakthrough-like occupations in order to avoid spiraling. We leave these questions to future research.

A Proofs for Section 4

Continuation of the proof of Lemma 4.1. We now show the inequality for the breakthrough environment. Let $q_a > q_b$. The employer uses worker a exclusively for a period of length $t^* = \frac{1}{\lambda_h} \log \frac{q_a(1-q_b)}{(1-q_a)q_b}$ and then splits the task equally among the two workers for a subsequent period of length $t_s := \frac{2}{\lambda_h} \log \frac{q_b(1-\underline{p})}{(1-q_b)\underline{p}}$. Let $S(h, q_b)$ and $S(\ell, q_b)$ denote the payoffs to a high-type worker and a low-type worker, respectively, if (i) his opponent is a high type with probability q_b; (ii) the employer holds the same belief about both workers and hence splits the task equally between the two workers until the belief drops to $\underline{p}$. The post-investment payoff for each worker and type are:

$$
\begin{aligned}
U_a(h; q_a, q_b) &= 1 - e^{-rt^*} + e^{-rt^*} \left(1 - e^{-\lambda_h t^*} + e^{-\lambda_h t^*} S(h, q_b)\right), \\
U_a(\ell; q_a, q_b) &= 1 - e^{-rt^*} + e^{-rt^*} S(\ell, q_b), \\
U_b(h; q_a, q_b) &= e^{-rt^*} \left(1 - q_a + q_a e^{-\lambda_h t^*}\right) S(h, q_b), \\
U_b(\ell; q_a, q_b) &= e^{-rt^*} \left(1 - q_a + q_a e^{-\lambda_h t^*}\right) S(\ell, q_b).
\end{aligned}
$$

Note that $U_a(h; q_a, q_b) - U_a(\ell; q_a, q_b) > e^{-rt^*}(S(h, q_b) - S(\ell, q_b))$ whereas $U_b(h; q_a, q_b) - U_b(\ell; q_a, q_b) < e^{-rt^*}(S(h, q_b) - S(\ell, q_b))$. Hence, $B_a(q_a, q_b) > B_b(q_a, q_b)$.

To characterize $S(h, q_b)$ and $S(\ell, q_b)$, let t_1 be the arrival time of a high type's breakthrough and t_2 the arrival time of his opponent's breakthrough. For a low type, a breakthrough never arrives. In the absence of any breakthroughs, the employer experiments with the workers until the belief hits $\underline{p}$. The length of this experimentation period is given by t_s as defined above. The CDFs of t_1 and t_2 for $t_1, t_2 \leqslant t_s$ are:[19]

$$
F_1(t_1) = 1 - e^{-\frac{\lambda_h t_1}{2}}, \quad F_2(t_2) = q_b(1 - e^{-\frac{\lambda_h t_2}{2}}),
$$

with corresponding density functions f_1 and f_2 respectively. Therefore,

$$
S(\ell, q_b) = \int_0^{t_s} f_2(t_2) \frac{1 - e^{-rt_2}}{2} \, \mathrm{d}t_2 + (1 - F_2(t_s)) \frac{1 - e^{-rt_s}}{2},
$$

[19]When the task is split equally among workers, the arrival rate for each worker is $\lambda_h/2$.

$$S(h,q_b)=\int_0^{t_s} f_1(t_1)\left(\int_0^{t_1} f_2(t_2)\frac{1-e^{-rt_2}}{2}\,\mathrm{d}t_2+(1-F_2(t_1))\left(\frac{1-e^{-rt_1}}{2}+e^{-rt_1}\right)\right)\mathrm{d}t_1$$
$$+(1-F_1(t_s))\left(\int_0^{t_s} f_2(t_2)\frac{1-e^{-rt_2}}{2}\,\mathrm{d}t_2+(1-F_2(t_s))\frac{1-e^{-rt_s}}{2}\right).$$

This allows us to obtain explicit expressions for B_a and B_b. Letting $\mu_h := \lambda_h/r$, we have

$$B_a(q_a,q_b)=\pi\left(\frac{q_b(\underline{p}-1)}{(q_b-1)\underline{p}}\right)^{-2/\mu_h}\left(\frac{(q_b-1)q_a}{q_b(q_a-1)}\right)^{-1/\mu_h}$$
$$\cdot\frac{(\underline{p}-1)^2\left(\frac{q_b(\underline{p}-1)}{(q_b-1)\underline{p}}\right)^{2/\mu_h}(q_b(\mu_h q_b+2)-(\mu_h+2)q_a)-(q_b-1)^2(\underline{p}(\mu_h(\underline{p}-2)-2)+(\mu_h+2)q_a)}{2(\mu_h+2)(q_b-1)(\underline{p}-1)^2 q_a}$$

if $q_a > q_b$, and

$$B_a(q_a,q_b)=\pi q_b\left(\frac{q_b(q_a-1)}{(q_b-1)q_a}\right)^{-\frac{\mu_h+1}{\mu_h}}\left(\frac{(\underline{p}-1)q_a}{\underline{p}(q_a-1)}\right)^{-2/\mu_h}$$
$$\cdot\frac{\mu_h(\underline{p}-1)^2 q_a\left(\frac{(\underline{p}-1)q_a}{\underline{p}(q_a-1)}\right)^{2/\mu_h}-(q_a-1)(\underline{p}(\mu_h(\underline{p}-2)-2)+(\mu_h+2)q_a)}{2(\mu_h+2)(\underline{p}-1)^2 q_a^2}$$

if $q_a \leqslant q_b$. At any (q_a,q_b) such that $q_a \neq q_b$, it is immediate that B_a is continuously differentiable. Moreover,

$$\lim_{q_a\to q_b^+} B_a(q_a,q_b)=\lim_{q_a\to q_b^-} B_a(q_a,q_b)$$
$$\lim_{q_a\to q_b^+}\frac{\partial B_a(q_a,q_b)}{\partial q_a}=\lim_{q_a\to q_b^-}\frac{\partial B_a(q_a,q_b)}{\partial q_a}$$
$$\lim_{q_a\to q_b^+}\frac{\partial B_a(q_a,q_b)}{\partial q_b}=\lim_{q_a\to q_b^-}\frac{\partial B_a(q_a,q_b)}{\partial q_b}.$$

Hence, B_a is continuously differentiable at $q_a = q_b$ as well.[20]

■

Proof of Proposition 4.1. A belief pair (q_a,q_b) and a cost-threshold pair (c_a,c_b) consist in an

[20]For detailed calculations, see this online supplement.

equilibrium if and only if $\forall i \in \{a, b\}$:

$$B_i(q_a, q_b) = c_i, \text{ and } q_i = p_i + (1 - p_i)F(c_i)\pi.$$

From the second condition, we have $c_i = F^{-1}\left(\frac{q_i - p_i}{(1-p_i)\pi}\right)$. Hence, a belief pair (q_a, q_b) consists an equilibrium if and only if:

$$\begin{cases} \dfrac{1}{\pi} B_a(q_a, q_b) - \dfrac{1}{\pi} F^{-1}\left(\dfrac{q_a - p_a}{(1-p_a)\pi}\right) & = 0 \\ \dfrac{1}{\pi} B_b(q_a, q_b) - \dfrac{1}{\pi} F^{-1}\left(\dfrac{q_b - p_b}{(1-p_b)\pi}\right) & = 0. \end{cases} \tag{4}$$

Let $g_a(p_a, p_b, q_a, q_b)$ and $g_b(p_a, p_b, q_a, q_b)$ denote respectively the LHS of each equation. Both g_a and g_b are continuously differentiable, because B_a, B_b and F are continuously differentiable and F' is strictly positive.

Existence of symmetric equilibrium. We first show that if workers have the same prior belief, there is a symmetric equilibrium in which they have the same post-investment belief. Let $\hat{p}$ denote the two workers' prior belief. A symmetric equilibrium exists if there exists $\hat{q} \in [\hat{p}, \hat{p} + (1 - \hat{p})\pi]$ such that:

$$\begin{aligned} & B_i(\hat{q}, \hat{q}) = F^{-1}\left(\frac{\hat{q} - \hat{p}}{(1-\hat{p})\pi}\right) \implies \\ & \frac{\pi\left(\mu_h + \frac{\left(\frac{\hat{q}(1-\underline{p})}{(1-\hat{q})\underline{p}}\right)^{-\frac{\mu_h+2}{\mu_h}}((\mu_h+2)\hat{q}+\underline{p}(\mu_h(\underline{p}-2)-2))}{(1-\underline{p})\underline{p}}\right)}{2(\mu_h+2)} = F^{-1}\left(\frac{\hat{q} - \hat{p}}{\pi(1-\hat{p})}\right). \end{aligned} \tag{5}$$

Such a $\hat{q}$ exists because for $\hat{q} \in [\hat{p}, \hat{p} + (1 - \hat{p})\pi]$: (i) $B_i(\hat{q}, \hat{q})$ is continuous, strictly positive, and strictly less than one; and (ii) $F^{-1}\left(\frac{\hat{q}-\hat{p}}{(1-\hat{p})\pi}\right)$ is strictly increasing, equals 0 if $\hat{q} = \hat{p}$, and equals 1 if $\hat{q} = \hat{p} + (1 - \hat{p})\pi$. Therefore, there exists $\hat{q} \in (\hat{p}, \hat{p} + (1 - \hat{p})\pi)$ such that $F^{-1}\left(\frac{\hat{q}-\hat{p}}{(1-\hat{p})\pi}\right)$ crosses $B_i(\hat{q}, \hat{q})$ from below. Hence, $g_a(\hat{p}, \hat{p}, \hat{q}, \hat{q}) = g_b(\hat{p}, \hat{p}, \hat{q}, \hat{q}) = 0$.

Non-singularity of the Jacobian at $(\hat{p}, \hat{p}, \hat{q}, \hat{q})$. We next show that the Jacobian matrix evaluated at $(\hat{p}, \hat{p}, \hat{q}, \hat{q})$ is invertible for a generic set of parameters, where the Jacobian is given by:

$$J = \left(\begin{matrix} \frac{\partial g_a}{\partial q_a} & \frac{\partial g_a}{\partial q_b} \\ \frac{\partial g_b}{\partial q_a} & \frac{\partial g_b}{\partial q_b} \end{matrix}\right)\Bigg|_{(\hat{p},\hat{p},\hat{q},\hat{q})}.$$

Note that J is symmetric: $\frac{\partial g_a}{\partial q_a} = \frac{\partial g_b}{\partial q_b}\Big|_{(\hat{p},\hat{p},\hat{q},\hat{q})}$ and $\frac{\partial g_a}{\partial q_b} = \frac{\partial g_b}{\partial q_a}\Big|_{(\hat{p},\hat{p},\hat{q},\hat{q})}$. Hence, we only need to show that:

$$\frac{\partial g_a}{\partial q_a} + \frac{\partial g_a}{\partial q_b}\bigg|_{(\hat{p},\hat{p},\hat{q},\hat{q})} \neq 0 \tag{6}$$

$$\frac{\partial g_a}{\partial q_a} - \frac{\partial g_a}{\partial q_b}\bigg|_{(\hat{p},\hat{p},\hat{q},\hat{q})} \neq 0. \tag{7}$$

Claim (6) holds because

$$\frac{\partial g_a}{\partial q_a} + \frac{\partial g_a}{\partial q_b}\bigg|_{(\hat{p},\hat{p},\hat{q},\hat{q})} = \left(\frac{\partial B_a(q_a, q_b)}{\partial q_a} + \frac{\partial B_a(q_a, q_b)}{\partial q_b} - \frac{\mathrm{d}\ F^{-1}\left(\frac{q_a - p_a}{(1-p_a)\pi}\right)}{\mathrm{d}\ q_a}\right)\Bigg|_{(\hat{p},\hat{p},\hat{q},\hat{q})} < 0.$$

The inequality follows from the fact that $\frac{1}{\pi}F^{-1}\left(\frac{q-p_a}{(1-p_a)\pi}\right)$ generically crosses $\frac{1}{\pi}B_a(q,q)$ transversally from below at $q = \hat{q}$, as shown in the lemma below.

Lemma A.1. *There exists* $\Pi \subset (0,1)$ *of measure one such that*

$$g(q,\pi) := \frac{1}{\pi}B_a(q,q) - \frac{1}{\pi}F^{-1}\left(\frac{q-\hat{p}}{(1-\hat{p})\pi}\right)$$

intersects zero transversally for any $\pi \in \Pi$.

Proof. First, $g(q,\pi)$ is strictly increasing in π because the term $\frac{1}{\pi}B_a(q,q)$ is independent of π and F^{-1} is strictly increasing in $[0,1]$. Therefore 0 is a regular value of $g(q,\pi)$. By the transversality theorem, there exists a set $\Pi \in (0,1)$ of values for π such that $(0,1)\setminus\Pi$ has measure zero and for any $\pi \in \Pi$, 0 is a regular value of $g(q,\pi)$. Hence, generically the derivative of $g(q,\pi)$ with respect to q at $q = \hat{q}$ is non-zero. ■

Claim (7) holds unless:

$$\frac{\frac{\left(\frac{\hat{q}(1-\underline{p})}{(1-\hat{q})\underline{p}}\right)^{-2/\mu_h}\left((\mu_h+2)\hat{q}^2+\mu_h(2\hat{q}-1)\underline{p}^2-2(\mu_h+1)(2\hat{q}-1)\underline{p}\right)}{(\underline{p}-1)^2}+\frac{2\hat{q}(\mu_h\hat{q}+1)}{1-\hat{q}}}{2(\mu_h+2)\hat{q}^2}=\frac{1}{\pi^2(1-\hat{p})F'\left(F^{-1}\left(\frac{\hat{q}-\hat{p}}{\pi(1-\hat{p})}\right)\right)}. \tag{8}$$

Fix $(F,\underline{p},\mu_h,\hat{p})$. The following lemma shows that for almost any $(\pi,\hat{p})$ claim (7) holds.

Lemma A.2. *Suppose that F is weakly convex. Then, claim (7) is satisfied in equilibrium for almost all $(\pi,\hat{p})$.*

Proof. Equations (5) and (8) are equivalent respectively to:

$$\begin{aligned}
g_1(p,\pi,q) &= \frac{1}{\pi}F^{-1}\left(\frac{q-p}{(1-p)\pi}\right)-h_1(q)=0\\
g_2(p,\pi,q) &= \frac{1}{\pi^2(1-p)F'\left(F^{-1}\left(\frac{q-p}{\pi(1-p)}\right)\right)}-h_2(q)=0,
\end{aligned}$$

where h_1,h_2 are functions of q only. The determinant of the Jacobian matrix of this system with respect to (p,π) is

$$\frac{(1-p)\pi F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)F'\left(F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)\right)+2-p-q}{(1-p)^3\pi^5F'\left(F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)\right)^2}+\frac{(1-q)F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)F''\left(F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)\right)}{(1-p)^3\pi^5F'\left(F^{(-1)}\left(\frac{p-q}{(p-1)\pi}\right)\right)^3},$$

which is strictly positive if F'' is weakly positive. So the Jacobian matrix is invertible.

This implies that for almost all π, the function $g=(g_1,g_2)(p,q)$ crosses $(0,0)$ transversally, which means the points at which (p,q) form a symmetric equilibrium (i.e., satisfying (5)) and violate claim (7) (i.e., satisfying (8)) are isolated. Hence, if F is weakly convex, for almost all π, the set of p for which the symmetric equilibrium violates claim (7) is finite. Hence, claim (7) is satisfied in equilibrium for almost all $(\pi,\hat{p})$. ■

Implicit function theorem. We can invoke the implicit function theorem everywhere except for a set of measure zero of parameters. Therefore, by the implicit function theorem, there exists a neighborhood $B \subset [0,1]^2$ of $(\hat{p},\hat{p})$ and a unique continuously differentiable map $\mathbf{q}: B \to [0,1]^2$ such that $g_a(\hat{p},\hat{p},\mathbf{q}(\hat{p},\hat{p})) = 0$, $g_b(\hat{p},\hat{p},\mathbf{q}(\hat{p},\hat{p})) = 0$ and for any $(p_a,p_b) \in B$

$$g_a(p_a,p_b,\mathbf{q}(p_a,p_b)) = g_b(p_a,p_b,\mathbf{q}(p_a,p_b)) = 0.$$

By the continuity of the map $\mathbf{q}$, $\mathbf{q}(p_a,p_b)$ converges to $\mathbf{q}(\hat{p},\hat{p}) = (\hat{q},\hat{q})$ as $p_a \to \hat{p}$ and $p_b \to \hat{p}$. Hence, the workers' post-investment probabilities of being a high type converge as well.

■

Proof for Proposition 4.2. Throughout the proof, "worker's type" refers to the worker's pre-investment type. We focus on the equilibrium with post-investment beliefs $q_a > q_b$ and cost thresholds $c_a > c_b$ as $p_b \uparrow p_a$. The argument for the equilibrium in which $q_b > q_a$ is similar.

We first characterize this equilibrium. Using B_a and B_b derived in the proof of Lemma 4.1, the cost thresholds are:

$$c_a = \pi \frac{\mu_\ell}{\mu_\ell + 1} > c_b = \pi \frac{\mu_\ell^2 (1-q_a)}{(\mu_\ell+1)^2}.$$

where the post investment belief pair (q_a,q_b) is given by $q_a = p_a + (1-p_a)\pi F(c_a)$ and $q_b = p_b + (1-p_b)\pi F(c_b)$. Note that $c_i \in (0,1)$ for each $i \in \{a,b\}$. Given that $c_a > c_b$ and $p_a > p_b$, the employer is indeed willing to favor worker a.

Let $\kappa := \frac{\mu_\ell(1-q_a)}{\mu_\ell+1} < 1$. Since worker a is favored, a high-type worker a obtains payoff 1, while a high-type worker b obtains payoff κ. Hence, the ratio of worker b's to worker a's payoff, conditional on each being a high type, is κ.

We next argue that for any realized cost c, a low-type worker b's payoff is *at most* a fraction κ of the low-type worker a's payoff. Hence, the same holds when taking the expectation with respect to c.

1. If $c \geqslant c_a$, neither low-type worker a nor low-type worker b invests. The ratio of low-type worker b's payoff to low-type worker a's payoff is exactly κ.

2. If $c_b < c < c_a$, a low-type worker a is willing to invest but a low-type worker b is not. If the low-type worker a deviates to no investment, the ratio of low-type worker

b's payoff to low-type worker a's payoff is κ. By investing worker a obtains a strictly higher payoff. Therefore, the payoff ratio must be strictly lower when the low-type worker a invests.

3. If $c \leqslant c_b$, both the low type of worker a and of worker b invest. Ignoring investment cost $c > 0$, the payoff ratio of the low-type worker b to that of the low-type worker a is κ. Once the investment cost is subtracted from both the numerator and the denominator, the payoff ratio becomes strictly smaller.

■

Proof of Proposition 4.3. Throughout the proof, we set $\pi = 1$ without loss, as π merely scales the benefit from investment $B_i(q_a, q_b)$ and i's threshold for investment for each i. Let i denote the post-investment favored worker, and $-i$ be the discriminated one.

As we take λ_ℓ, λ_h to infinity, worker i's benefit from investment converges to 1 under breakdown learning, while it converges to

$$\bar{B}_i(q_i, q_{-i}, \underline{p}) := \frac{\frac{(1-q_{-i})^2\left(q_i+\underline{p}^2-2\underline{p}\right)}{(1-\underline{p})^2} + q_i - q_{-i}^2}{2q_i(1-q_{-i})},$$

under breakthrough learning. This function $\bar{B}_i(q_i, q_{-i}, \underline{p})$ increases in q_i, and decreases in q_{-i} and $\underline{p}$. Since q_i is bounded above by $\bar{p}_a := p_a + (1-p_a)\pi$, q_{-i} is bounded below by p_b, and $\underline{p}$ is bounded below by 0, $\bar{B}_i(q_i, q_{-i}, \underline{p})$ is bounded from above by $\bar{B}_i(\bar{p}_a, p_b, 0) = \frac{{p_b}^2-\bar{p}_a((p_b-2)p_b+2)}{2\bar{p}_a(p_b-1)}$, which is strictly smaller than 1. By continuity, when λ_ℓ, λ_h are sufficiently large, the post-investment favored worker invests more under breakdown learning than under breakthrough learning.

As we take λ_ℓ, λ_h to infinity, worker $-i$'s benefit from investment converges to $(1-q_i)$ under breakdown learning, while it converges to

$$\bar{B}_{-i}(q_i, q_{-i}, \underline{p}) := \frac{(1-q_i)\left((2-q_{-i})q_{-i}+\underline{p}^2-2\underline{p}\right)}{2(1-q_{-i})q_{-i}(1-\underline{p})^2},$$

under breakthrough learning. The function $\bar{B}_{-i}(q_i, q_{-i}, \underline{p})$ increases in q_{-i}, and decreases in

q_i and $\underline{p}$. As $\underline{p}$ converges to zero, the benefit converges to

$$\bar{B}_{-i}(q_i, q_{-i}, 0) = \frac{(1-q_i)(2-q_{-i})}{(2-2q_{-i})} > 1 - q_i.$$

Here, the inequality follows from $0 < q_{-i} < 1$. Given that the favored worker i invests more under breakdown than under breakthrough learning, q_i is higher under breakdown as well. Hence, the discriminated worker's benefit from investment is higher under breakthrough than under breakdown when $\underline{p}$ is small enough and λ_h, λ_ℓ are large enough. ■

B Supplementary material for Section 5

This section extends our baseline model to a two-sided matching market with a continuum of workers and employers. There is a unit mass of employers, a mass of size $\alpha > 1$ of a-workers, and a mass of size $\beta > 0$ of b-workers. Employers are ex ante homogeneous. Both employers and workers are long-lived. They share the same discount rate $r > 0$. At $t = 0$, each worker's type is drawn independently from other workers' types. An a-worker's type is high with probability p_a, and a b-worker's type is high with probability p_b. There is also a unit mass of identical safe arms available. We assume that $p_a > p_b > \underline{p}$.

Both employers and workers are long-lived. They share the same discount rate $r > 0$. At each instant, each employer has one task to allocate and each worker can take up at most one task.

Matching protocol. At each $t \geqslant 0$, the following frictionless matching protocol takes place:

(i) each unmatched employer is matched randomly with an unmatched worker;

(ii) the matched employer and worker decide simultaneously whether to accept the match;

(iii) if at least one rejects, they return to the unmatched pool and are rematched instantaneously;

(iv) if an employer rejects all unmatched workers, she takes a safe arm;

(v) this process ends when each employer is either matched to a worker or takes the safe arm.

An employer returns to the unmatched pool if she fires the worker that she is currently matched to. We assume that all signals are observed publicly. Therefore, all employers hold the same belief about the type of any worker at any time t. The frictionless matching protocol ensures that at each instance the most productive workers are matched provided that their probabilities of being a high type are higher than $\underline{p}$.

Task scarcity. When there are more workers than tasks (i.e., $\alpha + \beta > 1$), not all workers are matched immediately at $t = 0$. Such relative scarcity of tasks is both necessary and sufficient for spiraling to arise in breakdown learning. To simplify exposition, we impose a stronger assumption: $\alpha > 1$.[21] This assumption guarantees that only a-workers are matched at $t = 0$.

B.1 Breakthrough learning

Once a worker generates a breakthrough, his employer keeps him for the rest of time. To track how many workers have "secured their jobs", we let $\omega(t) \in [0, 1]$ be the mass of workers who have generated a breakthrough by t, so $(1 - \omega(t))$ is the mass of employers who are still learning about the type of their current match.

At $t = 0$, all employers are matched to a-workers due to $\alpha > 1$. Within the next instant, the belief for those matched a-workers who have not generated a breakthrough drops slightly below p_a. Their employers find it optimal to switch to previously unmatched a-workers, the belief for whom is p_a. This is essentially equivalent to all a-workers being employed and allocated $1/\alpha < 1$ of a task at $t = 0$.

In the next instant, those a-workers who have generated a breakthrough stay matched forever. Those who have not are once again allocated a fraction of a task. This process goes on until the belief for those a-workers without a breakthrough drops to p_b. We let T_b denote this time, which is deterministic. From T_b onward, employers start allocating tasks to b-workers as well. This T_b is the delay that is experienced by group b uniformly.

We let $q(t)$ denote the belief for a matched worker who has not generated a breakthrough until time t. For any $t \in [0, T_b)$, a mass $(\alpha - \omega(t))$ of a-workers have not generated a breakthrough. Each is a high type with probability $q(t)$, and is allocated $\frac{1-\omega(t)}{\alpha-\omega(t)} \in (0, 1)$ of a

[21] Our analysis extends to the case of $\alpha < 1$ and $\alpha + \beta > 1$.

task. Therefore, the evolution of $\omega(t)$ follows:

$$\mathrm{d}\omega(t) = (\alpha - \omega(t))q(t)\lambda_h \frac{1-\omega(t)}{\alpha - \omega(t)}\mathrm{d}t = q(t)\lambda_h(1-\omega(t))\mathrm{d}t \quad \text{and} \quad \omega(0) = 0. \tag{9}$$

By the law of large numbers, for any $t \in [0, T_b)$, $q(t)$ satisfies:

$$q(t)(\alpha - \omega(t)) + \omega(t) = p_a\alpha \quad \Longrightarrow \quad q(t) = \frac{\alpha p_a - \omega(t)}{\alpha - \omega(t)}. \tag{10}$$

The value T_b is given by $q(T_b) = p_b$.

Starting from T_b, employers who did not have a breakthrough over $[0, T_b)$ start allocating tasks over a larger set of workers: a-workers who have not generated a breakthrough until time T_b and all b-workers. The method for solving for $\omega(t), q(t)$ is similar. The evolution of $\omega(t)$ is the same as (9). By the law of large numbers, for any $t \geqslant T_b$, $q(t)$ satisfies:

$$q(t)(\alpha + \beta - \omega(t)) + \omega(t) = p_a\alpha + p_b\beta \quad \Longrightarrow \quad q(t) = \frac{\alpha p_a + \beta p_b - \omega(t)}{\alpha + \beta - \omega(t)}.$$

The process ends when either $\omega(t)$ reaches 1 or $q(t)$ reaches $\underline{p}$, depending on which event occurs earlier. If $\omega(t)$ reaches 1 first, then all employers are matched with workers who have generated a breakthrough. Otherwise, some employers take safe arms when $q(t)$ drops to $\underline{p}$.

Proposition B.1 (Self-correction under breakthroughs)**.** *For $\alpha > 1$ and $\beta > 0$, the expected payoff of an a-worker converges to that of a b-worker as $p_b \uparrow p_a$.*

Proof. We first show that as $p_b \uparrow p_a$, $T_b \to 0$. By the definition of T_b and the expression for $q(t)$ in (10), we have that

$$\omega(T_b) = \frac{\alpha(p_a - p_b)}{1 - p_b}.$$

Therefore, as $p_b \uparrow p_a$, $\omega(T_b) \to 0$. Using the fact that $\omega(0) = 0$, $\omega(t)$ is independent of p_b, and it is strictly increasing in t, we conclude that $T_b \downarrow 0$.

Conditional on reaching T_b without a breakthrough, an a-worker has the same continuation payoff as a b does. As $T_b \to 0$, the probability of a breakthrough over $[0, T_b)$ goes to zero and so does the flow payoff from being allocated the task over $[0, T_b)$. Hence, the payoff of an a-worker approaches that of a b-worker as $T_b \to 0$. ∎

B.2 Breakdown learning

Under breakdowns, a matched worker stays matched as long as no breakdown occurs. At time 0, a unit mass of a-workers are matched with employers. When a matched worker generates a breakdown, his employer replaces him with an a-worker who has never been tried before. This process goes on until all the a-workers are tried. From that moment on, an employer who just experienced a breakdown hires a b-worker who has never been tried before. We let T_b denote the first time that a b-worker is hired. This T_b is again the delay that is experienced by group b uniformly.

We let $\underline{\omega}(t) \geqslant 1$ be the mass of workers who have been tried before t. Among these workers, one unit are currently employed, and a mass $(\underline{\omega}(t) - 1)$ of workers have generated a breakdown before t.

For any $t \in [0, T_b)$, the mass of employers who are matched to high-type workers are $p_a\underline{\omega}(t)$, so $1 - p_a\underline{\omega}(t)$ are matched to low-type workers. Hence, the evolution of $\underline{\omega}(t)$ follows:

$$\mathrm{d}\underline{\omega}(t) = (1 - p_a\underline{\omega}(t))\lambda_\ell \mathrm{d}t.$$

This along with the boundary condition $\underline{\omega}(0) = 1$ pins down $\underline{\omega}(t)$ for any $t \in [0, T_b)$:

$$\underline{\omega}(t) = \frac{1 - (1 - p_a)e^{-\lambda_\ell p_a t}}{p_a}.$$

If $p_a\alpha < 1$, then T_b is finite and solves $\underline{\omega}(T_b) = \alpha$. Otherwise T_b is infinity.

Suppose that $p_a\alpha < 1$. For any $t \geqslant T_b$, the mass of employers who are matched to high-type workers are $p_a\alpha + p_b(\underline{\omega}(t) - \alpha)$. Hence, the evolution of $\underline{\omega}(t)$ follows:

$$\mathrm{d}\underline{\omega}(t) = (1 - p_a\alpha - p_b(\underline{\omega}(t) - \alpha))\lambda_\ell \mathrm{d}t.$$

This along with the boundary condition $\underline{\omega}(T_b) = \alpha$ pins down $\underline{\omega}(t)$ for any $t \geqslant T_b$:

$$\underline{\omega}(t) = \frac{1 - (1 - \alpha p_a)e^{\lambda_\ell p_b (T_b - t)} - \alpha(p_a - p_b)}{p_b}$$

This process of hiring untried b-workers ends when either $\underline{\omega}(t)$ reaches $\alpha+\beta$ or $p_a\alpha+p_b(\underline{\omega}(t) - \alpha)$ reaches 1, depending on which event occurs earlier. We let T_s denote the time at which this process ends. If $\underline{\omega}(t)$ reaches $\alpha + \beta$ first, then at time T_s all workers have been tried. Otherwise, some b-workers are never matched to any employer because all employers are already matched to high-type workers.

Proposition B.2 (Spiraling under breakdowns)**.** *As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker is strictly less than one.*

Proof. Suppose first that $\alpha p_a > 1$. A b-worker's payoff is zero, so the ratio is zero as well. The statement holds trivially.

Next, let $1 < \alpha < 1/p_a$. This assumption guarantees that $0 < T_b < \infty$. Let $V(p_i)$ denote a *worker's continuation payoff from the time he is first allocated the task.* From the proof of Proposition 3.2, we know that $V(p_i) = p_i + (1 - p_i)r/(\lambda_\ell + r)$. An a-worker's expected payoff is

$$\frac{1}{\alpha}\left(V(p_a) + \int_0^{T_b} e^{-rt}V(p_a)\, \mathrm{d}\underline{\omega}(t)\right).$$

A b-worker's expected payoff is

$$\frac{1}{\beta}\int_{T_b}^{T_s} e^{-rt}V(p_b)\, \mathrm{d}\underline{\omega}(t).$$

As $p_b \uparrow p_a$, $V(p_b) \uparrow V(p_a)$. But because each b-worker gets a chance strictly later than any a-worker, a b-worker's expected payoff is strictly lower than that of an a-worker. ■

Spiraling arises if and only if b-workers are not guaranteed to be allocated the task at time $t = 0$. That is, tasks must be relatively scarce. For simplicity, we assumed that $\alpha > 1$ so that b-workers never get a chance at $t = 0$. But even if some b-workers get a chance at $t = 0$, the expected payoffs of the two groups do not converge as $p_b \uparrow p_a$ for as long as other b-workers are delayed. Proposition 5.1 shows that the larger is the labor force — i.e., the larger is the mass of workers relative to the unit mass of tasks — the greater is the inequality across groups.

Proof for Proposition 5.1. The rest of this argument supposes that $p_a(\alpha + \beta) < 1$. The argument for $p_a(\alpha + \beta) \geqslant 1$ is similar, and hence omitted.

Using the expression we have for $\underline{\omega}$ and applying the change of variables $\mu_\ell = \lambda_\ell/r$, we compute the expected payoffs of workers from each group. The ratio of the expected payoff of an a-worker to that of a b-worker is:

$$-\frac{\beta(\mu_\ell p_b+1)\left((\mu_\ell+1)\left(\frac{p_a-1}{\alpha p_a-1}\right)^{\frac{1}{\mu_\ell p_a}}+\mu_\ell(\alpha p_a-1)\right)\left(\frac{\alpha p_a-1}{\alpha p_a+\beta p_b-1}\right)^{\frac{1}{\mu_\ell p_b}}}{\alpha\mu_\ell(\mu_\ell p_a+1)\left((\alpha p_a-1)\left(\frac{\alpha p_a-1}{\alpha p_a+\beta p_b-1}\right)^{\frac{1}{\mu_\ell p_b}}-\alpha p_a-\beta p_b+1\right)}.$$

We take the limit of this ratio as $p_b \uparrow p_a$ and differentiate with respect to α and β. By applying the change of variables $z=\frac{1-p_a}{1-\alpha p_a}>1$ and $y=\frac{1-\alpha p_a}{1-p_a(\alpha+\beta)}>1$ to replace α and β and simplify the algebra, it follows that these two derivatives are both positive. ■

C Supplementary material for Section 6

C.1 Setup

This section introduces endogenous, flexible wages in the context of the two-sided market introduced in Section 5 and Appendix B. We show that the self-correcting property of breakthroughs and the spiraling property of breakdowns still hold. In particular, wage flexibility is insufficient to prevent spiraling under breakdowns.

In this section, we use $i \in [0, \alpha+\beta]$ to index a worker. Worker i is from group a if $i \in [0,\alpha]$ and from group b if $i \in (\alpha, \alpha+\beta]$. We use $j \in [0,1]$ to index an employer. At each instant, each employer demands one unit of labor and each worker supplies one unit of labor. There is also a unit mass of identical safe arms. An employer takes a safe arm whenever not matched to a worker. We assume that no one observes the workers' realized types and all learning is public.

Stage-game matching. We consider one-to-one matching between workers and employers. A stage-game matching describes how workers are matched to employers along with a wage for each matched pair. We assume that workers are protected by limited liability, so wages are nonnegative.

Let $D_{ij} \in \{1,0\}$ indicate whether worker i and employer j are matched to each other, and if they are let $W_{ij} \geqslant 0$ denote the wage. If $D_{ij}=1$, worker i's payoff is W_{ij} and employer j's payoff is $p_i v - W_{ij}$, where p_i denotes the probability that worker i's type is high. If $D_{ij}=0$ for all j, worker i is unmatched and gets zero payoff. If $D_{ij}=0$ for all i, employer j takes a safe arm and gets a fixed payoff of $s>0$, which corresponds to a belief threshold $p_s := s/v$.

We assume that $p_s < p_b$. Let $\mathcal{D}$ be the set of all stage-game matchings.

Dynamic matching. Let $\mathcal{H} := \cup_{t \geqslant 0} \mathcal{H}_t$ be the set of all histories and $\mathcal{H}_t$ the set of all time-t histories. All signals are publicly observed, hence a time-t history consists of all past matchings and realized signals until t. A *dynamic matching* $\mu = (\mu_t)_{t \geqslant 0}$ specifies a lottery over stage-game matchings for any history, i.e., $\mu_t : \mathcal{H}_t \to \Delta(\mathcal{D})$ for each t.

Solution concept. We first adopt the solution concept in Shapley and Shubik (1971) and define stable stage-game matchings. In Appendix C.2 we characterize the set of such matchings. Given a stage-game matching (D, W), (i, j) is called a *blocking pair* if they strictly prefer to be matched to each other at some wage $w \geqslant 0$ rather than following (D, W).

Definition 1. A stage-game matching (D, W) is *stable* if

(i) there exists no employer j who is matched to some i such that j strictly prefers to take a safe arm instead;

(ii) there exists no blocking pair.

Next we define dynamic stability based on the notion of a *stable convention* in Ali and Liu (2020).[22] In Appendix C.3 we show that repeating a stable stage-game matching is dynamically stable in both learning environments. For a given dynamic matching μ, let $\mu|_h$ denote the continuation matching after some history h.

Definition 2. A dynamic matching μ is *dynamically stable* if at every t and every history $h_t \in \mathcal{H}_t$, there exists no $\mathrm{d}t > 0$, however small, and

(i) no matched employer j under $\mu|_{h_t}$ who strictly prefers to take a safe arm over $[t, t + \mathrm{d}t)$ and then revert to $\mu|_{h_{t+\mathrm{d}t}}$;

(ii) no worker-employer pair (i, j) who strictly prefer to be matched to each other at some wage $w \geqslant 0$ over $[t, t + \mathrm{d}t)$ and then revert to $\mu|_{h_{t+\mathrm{d}t}}$;

(iii) no matched worker i under $\mu|_{h_t}$ who strictly prefers to be unmatched over $[t, t + \mathrm{d}t)$ and then revert to $\mu|_{h_{t+\mathrm{d}t}}$.

[22] Even though not crucial to our results, we assume that deviation wages are perfectly observable to all.

C.2 Stable stage-game matching

Recall that p_i denotes the probability that worker i's type is high. Let G denote the CDF of the distribution of p_i for $i \in [0, \alpha + \beta]$. Hence, $(\alpha + \beta)G(p)$ is the mass of workers with $p_i \leqslant p$. At time 0, p_i is either p_a or p_b, so $G(p)$ equals 0 if $p < p_b$, $\frac{\beta}{\alpha+\beta}$ if $p_b \leqslant p < p_a$, and 1 if $p \geqslant p_a$. As workers are matched to employers so more is learned about their types, G evolves over time.

In this subsection, we characterize the set of stable stage-game matchings for a fixed G. There exists a unique marginal belief p^M such that worker i is matched if $p_i > p^M$ and unmatched if $p_i < p^M$. Worker i's wage is a linear function of p_i.

Lemma C.1 (Equal profit across employers and linear wage)**.** *In any stable stage-game matching,*

1. *all employers make the same profit. If some employers take safe arms, then this profit is s;*

2. *if worker i is matched, his wage takes the form of $p_i v + c_1$, where c_1 is a constant.*

Proof. We first prove that employers make the same profit across all worker-employer pairs. Suppose that workers i_1 and i_2 are matched to employers j_1 and j_2 at wages w_1 and w_2 respectively. Let p_1 and p_2 be, respectively, the probabilities that i_1 and i_2 are high types. Suppose that employer j_1 makes a strictly higher profit than j_2:

$$vp_1 - w_1 > vp_2 - w_2.$$

Worker i_1 and employer j_2 can form a blocking pair at wage $w_1 + \epsilon$. Worker i_1's payoff improves by ϵ. Employer j_2's profit improves to $vp_1 - w_1 - \epsilon > vp_2 - w_2$. Hence, employers must make the same profit across all worker-employer pairs. This implies that the wage for worker i must take the form of $p_i v + c_1$.

What remains to be shown is that if some employers take safe arms, then all employers make a profit of s. If an employer makes more than s, he must be matched to a worker. Then an employer who is currently taking a safe arm can form a blocking pair with this worker. ∎

Based on Lemma C.1, a stable stage-game matching is without loss characterized by $(d(p), w(p))$, where $d(p)$ specifies the fraction of workers with $p_i = p$ who are matched and $w(p) = vp + c_1$ is the wage.

We next show that employers are matched to the most productive workers, provided that these workers are better than safe arms. We need to discuss two cases, depending on whether there exists a unit mass of workers who are preferred to safe arms. In order to distinguish these two cases, we look at the unit mass of most productive workers, and let p^* correspond to the least productive worker in this mass.

Definition 3. Let p^* be the highest probability p such that the mass of workers with $p_i \geqslant p$ is greater than 1:

$$(\alpha+\beta)\int_{p^*}^{1} \mathrm{d}G(s) \geqslant 1, \quad \text{and} \quad (\alpha+\beta)\int_{p}^{1} \mathrm{d}G(s) < 1, \quad \forall p > p^*.$$

Lemma C.2 shows that worker i is matched if $p_i > \max\{p^*, p_s\}$ and unmatched if $p_i < \max\{p^*, p_s\}$. Therefore, we call $\max\{p^*, p_s\}$ the *marginal belief* and let p^M denote the marginal belief.

Lemma C.2 (Most productive workers are matched)**.**

1. *Suppose that $p^* > p_s$. Then $d(p)$ equals 1 if $p > p^*$, and 0 if $p < p^*$. If there is no atom at p^*, then $d(p^*)$ can take any value in $[0,1]$. If there is an atom at p^*, then $d(p^*)$ is given by:*
$$(1 - G(p^*))(\alpha+\beta) + d(p^*)(G(p^*) - G(p^*-))(\alpha+\beta) = 1.$$
2. *Suppose that $p^* \leqslant p_s$. Then $d(p)$ equals 1 if $p > p_s$, and 0 if $p < p_s$. Moreover, $d(p_s)$ can take any value in $[0,1]$ subject to:*
$$(1 - G(p_s))(\alpha+\beta) + d(p_s)(G(p_s) - G(p_s-))(\alpha+\beta) \leqslant 1.$$

Proof. We prove the first part in two steps.

1. If a less productive worker is matched, then a more productive worker must be matched. By way of contradiction, suppose that a p_1 worker is matched and a p_2 worker is not,

and that $p_1 < p_2$. The employer who is matched to the p_1 worker can form a blocking pair with the p_2 worker.

2. If $p^* > p_s$, no employer takes a safe arm. Suppose otherwise. Then there exists an unmatched worker i with $p_i \geqslant p^*$. Then an employer who is taking a safe arm can form a blocking pair with this worker i.

We now prove the second part. Suppose that a worker's probability of being a high type is $p > p_s$ and he is unmatched. The mass of workers whose p_i are weakly above p is strictly smaller than 1. Hence, there exists an employer who is either matched to a worker with $p_i < p$ or taking a safe arm. This employer forms a blocking pair with the unmatched worker p. ■

We now fully characterize the wage function for matched workers. If $p^* > p_s$, we must distinguish two cases depending on whether there exists an unmatched worker whose belief is arbitrarily close to p^*. If such a worker exists, then the wage function is pinned down uniquely. Otherwise, if there is a belief gap between the least productive matched worker and the most productive unmatched worker, wage can take a range of values. If $p^* \leqslant p_s$, there always exists a safe arm for an employer to take, so the wage function is pinned down uniquely. Whenever unique, the wage for worker i is $(p_i - p^M)v$.

Lemma C.3 (Wage in stable stage-game matchings)**.**

1. *Suppose that $p^* > p_s$.*

   (1.a) If for any $\epsilon > 0$,
   $$\int_{p^*-\epsilon}^{p^*} (1 - d(s))\, dG(s) > 0,$$
   then $c_1 = -vp^*$ so $w(p_i) = (p_i - p^*)v$.

   (1.b) Otherwise, let p^{**} be the supremum belief among workers and safe arms whose belief is strictly smaller than p^*. Then the constant c_1 in $w(p_i) = vp_i + c_1$ can take any value in $[-vp^*, -vp^{**}]$.

2. *Suppose that $p^* \leqslant p_s$. Then $w(p_i) = (p_i - p_s)v$.*

Proof. We begin with showing that the wage function must be $w(p_i) = v(p_i - p^*)$ in the case of (1.a). First, the wage cannot be lower than this because of limited liability. Second, if $w(p^*) > 0$, then the employer that is matched to p^* worker can form a blocking pair with an unmatched worker whose p_i is arbitrarily close to p^*.

We now show (1.b). If there exists $\epsilon > 0$ such that

$$\int_{p^*-\epsilon}^{p^*} (1 - d(s))\mathrm{d}G(s) = 0,$$

then it must be that the fraction of workers whose belief is weakly above p^* is exactly 1. We argue that the constant c_1 in $w(p_i) = vp_i + c_1$ can be anything in:

$$c_1 \in [-vp^*, -vp^{**}].$$

Pick any c_1 in this range. All the employers get the same profit. Hence, an employer cannot form a blocking pair with another worker that is hired, since to attract that worker the employer has to offer a higher wage than $vp_i + c_1$. This will lead to a lower profit for the employer. Also, the employer cannot form a blocking pair with a worker that is not hired. The most profit the employer can make is vp^{**}, which is smaller than his current profit.

For the case of $p^* \leqslant p_s$, the proof is similar to that for the case of (1.a), so is omitted. ■

C.3 Stable dynamic matching

So far we characterized the set of stable stage-game matchings for any given G. The CDF G summarizes how much information there is about workers' types. Our characterization delineates how this information shapes workers' and employers' payoffs. In the dynamic setting, G evolves endogenously over time due to learning about workers' types. This section shows that repeating a stable stage-game matching after any history is dynamically stable.

Lemmata C.2 and C.3 showed that for certain G's there exist multiple stable stage-game matchings. Whenever such multiplicity arises, we select a stable stage-game matching that (i) leaves unmatched marginal-belief workers as much as possible, and (ii) assigns the employer-preferred wage. Such multiplicity arises on path very infrequently — in fact, wage multiplicity arises only at one point in time. Moreover, the selection criterion that we adopt is for ease of exposition only: the propositions below hold even with a different selection.

Definition 4. Fix G. Let $p^M(G)$ denote the marginal belief. Let $(d^*(p|G), w^*(p|G))$ be a stable stage-game matching that satisfies the following conditions:

1. $d^*(p^M(G)|G) = 0$ if $d(\cdot)$ is multi-valued at $p = p^M(G)$ in Lemma C.2;
2. $w^*(p|G) = (p - p^M(G))v$ is the employer-preferred wage function if $w(\cdot)$ is multi-valued in Lemma C.3.

Pick any history $h \in \mathcal{H}$. Let $G(h)$ denote the CDF of the distribution of p_i after history h. Let μ^* be the matching that always assigns the stable stage-game matching $(d^*(\cdot|G(h)), w^*(\cdot|G(h)))$ after every history h.

Proposition C.1. *Under either breakthrough or breakdown learning, μ^* is dynamically stable.*

Proof. Pick any $h_t \in \mathcal{H}_t$. We want to show that conditions (i)-(iii) in Definition 2 are satisfied in each learning environment.

(i) If employer j is matched to a worker under $\mu^*|_{h_t}$, his flow payoff on path is at least s. The distribution $G(h_{t+\mathrm{d}t})$, and hence j's continuation payoff from $t + \mathrm{d}t$ on, does not depend on j's deviation. Hence, he does not strictly prefer to take a safe arm over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*|_{h_{t+\mathrm{d}t}}$ in either learning environment.

(ii) Suppose that worker i and employer j are not matched to each other under $\mu^*|_{h_t}$. We next show that there is no wage $w \geqslant 0$ such that both i and j strictly prefer to be matched to each other at flow wage w over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*|_{h_{t+\mathrm{d}t}}$ in either learning environment.

If i is matched to another employer under $\mu^*|_{h_t}$, w needs to be strictly higher than worker i's current wage. This implies that employer j's flow payoff will be strictly lower than his current flow payoff. Hence, j does not strictly prefer to pair with i over $[t, t + \mathrm{d}t)$.

If i is not matched, this means that $p_i \leqslant p^M(G(h_t))$. But employer j's flow payoff on path is at least $p^M(G(h_t))v$. So employer j will not find it strictly profitable to be matched to i.

(iii) Suppose that worker i is matched at history h_t according to μ^*. Let $p(t)$ be this worker's probability of being a high type at history h_t. We next show that he does not strictly prefer to stay unmatched for $[t, t+\mathrm{d}t)$ and then revert to $\mu^*|_{h_{t+\mathrm{d}t}}$.

(a) We first consider breakdown learning. Pick any $\tau \geqslant t+\mathrm{d}t$. Let $Q(\tau)$ denote the probability that this worker has generated a breakdown in $[t,\tau)$, and $p(\tau)$ denote the probability that this worker is a high type at time τ conditional on no breakdown in $[t,\tau)$. By Bayes rule,

$$(1-Q(\tau))p(\tau) = p(t).$$

The worker's expected flow payoff at time τ is

$$\max\left\{0, (1-Q(\tau))\left(p(\tau)-p^M(G(h_\tau))\right)\right\} = \max\left\{0, p(t)\left(1-\frac{p^M(G(h_\tau))}{p(\tau)}\right)\right\} \tag{11}$$

which is weakly increasing in $p(\tau)$. Staying unmatched over $[t,t+\mathrm{d}t)$ and then reverting to $\mu^*|_{h_{t+\mathrm{d}t}}$ only makes $p(\tau)$ lower than its value on path, so the worker will not reject the match.

(b) We next consider breakthrough learning. Pick any $\tau \geqslant t+\mathrm{d}t$. Let $\tilde{Q}(\tau)$ denote the probability that this worker has generated a breakthrough in $[t,\tau)$, and $p(\tau)$ denote the probability that this worker is a high type at time τ conditional on no breakthrough in $[t,\tau)$. By Bayes rule,

$$\tilde{Q}(\tau) + (1-\tilde{Q}(\tau))p(\tau) = p(t).$$

The worker's expected flow payoff at time τ is

$$\begin{aligned}\tilde{Q}(\tau)(1-p^M(G(h_\tau)))v + (1-\tilde{Q}(\tau))\max\left\{0,(p(\tau)-p^M(G(h_\tau)))v\right\}\\ = \max\left\{\tilde{Q}(\tau)(1-p^M(G(h_\tau)))v, (p(t)-p^M(G(h_\tau)))v\right\}\end{aligned}$$

which is weakly increasing in $\tilde{Q}(\tau)$. Staying unmatched over $[t,t+\mathrm{d}t)$ and then reverting to $\mu^*|_{h_{t+\mathrm{d}t}}$ only makes $\tilde{Q}(\tau)$ lower than its value on path, so the worker will not reject the match.

■

Limited liability is not only sufficient, but also necessary for μ^* to be stable. If the wage can drop below zero, then workers have an incentive to be matched at negative wages in order to speed up learning about their type. Intuitively, the flow payoff to a worker of belief p_i is $\max\left\{v\left(p_i - p^M(G(h_t))\right), 0\right\}$ after history h_t. This flow payoff is convex in p_i. Hence, splitting a worker's prior belief strictly benefits this worker.

Our last proposition shows that the contrast between breakthrough and breakdown environments in terms of group inequality continues to hold. In particular, flexible wage does not close the payoff gap between group a and b in the breakdown environment.

Proposition C.2. *Given matching μ^*, as $p_b \uparrow p_a$ the expected payoff of an a-worker converges to that of a b-worker under breakthroughs but not under breakdowns.*

Proof. Let T_b be as defined in Appendix B.

Consider first the breakthrough environment. Because $\alpha > 1$, for an initial period $t \in [0, T_b)$, only a-workers are matched. If an a-worker has not achieved a breakthrough by T_b, his probability of being a high type is p_b. In this case, he has the same continuation payoff as a b-worker does. As $p_b \uparrow p_a$, $T_b \to 0$. Hence, worker a's payoff advantage vanishes as well.

We now consider the breakdown environment. Equation (11) in the proof of Proposition C.1 established that a worker who has been matched for longer has a higher expected flow payoff than a worker who has been matched for a shorter period. Hence, at any t the expected flow payoff of an a-worker is strictly higher than that of a b-worker. Moreover, group delay T_b does not converge to zero as $p_b \uparrow p_a$, hence an a-worker's payoff advantage due to $[0, T_b)$ does not converge to zero either. Hence, an a-worker's expected payoff is strictly higher than a b-worker's. ■

C.4 Wage gap under breakdown learning

We let $W_a(\tau)$ (resp., $W_b(\tau)$) denote the expected wage of a representative a-worker (resp., a representative b-worker) at any fixed time $\tau \geqslant 0$. To simplify exposition, we assume that (i) $\alpha > 1$, (ii) $\alpha p_a < 1$, and (iii) $\alpha p_a + \beta p_b > 1$. The first two conditions ensure that the delay for group b is positive but finite, i.e., $0 < T_b < \infty$. The third condition ensures that the pool of new workers is not exhausted before all employers identify a high-type worker. That is, there are more high-type workers than employers available. At the end of this section, we discuss the case of $\alpha p_a + \beta p_b \leqslant 1$.

We first solve for the expected flow wage at time τ of an i-worker first matched at time $t \leqslant \tau$. From expression (11), this expected flow wage is given by

$$p_i \left(1 - \frac{p^M(\tau)}{q(p_i, \tau - t)}\right),$$

where p_i is the prior belief of an i-worker, $p^M(\tau)$ is the marginal belief at time τ, and $q(p_i, \tau - t)$ is the employer's belief at time τ about an i-worker first matched at time $t \leqslant \tau$ who has not generated a breakdown over $[t, \tau)$. The marginal belief $p^M(\tau)$ is given by

$$p^M(t) = \begin{cases} p_a & \text{if } t \leqslant T_b \\ p_b & \text{otherwise,} \end{cases}$$

where the delay for group b is $T_b = \frac{1}{\lambda_\ell p_a} \log\left(\frac{1-p_a}{1-\alpha p_a}\right)$. Moreover,

$$q(p_i, \tau - t) = \frac{p}{p_i + (1 - p_i)e^{-\lambda_\ell(\tau - t)}}.$$

In order to calculate the expected wage of a representative i-worker, we also need the density over the time at which this i-worker is first matched. From Appendix B.2, we have the expression for $\underline{\omega}(t)$, the mass of workers who have been tried until time t:

$$\underline{\omega}(t) = \begin{cases} \dfrac{1 - (1 - p_a)e^{-\lambda_\ell p_a t}}{p_a} & \text{if } t \leqslant T_b \\ \dfrac{1 - (1 - \alpha p_a)e^{\lambda_\ell p_b (T_b - t)} - \alpha(p_a - p_b)}{p_b} & \text{otherwise.} \end{cases},$$

A unit mass of a-workers are matched at time 0. For any $t \in (0, T_b)$, new a-workers are tried at rate $\underline{\omega}'(t)$. For any $t \geqslant T_b$, new b-workers are tried at rate $\underline{\omega}'(t)$. Therefore, for any $\tau \geqslant 0$, the expected wage of a representative a-worker is

$$\frac{1}{\alpha}\left(p_a\left(1 - \frac{p^M(\tau)}{q(p_a, \tau)}\right) + \int_0^{T_b \wedge \tau} p_a \left(1 - \frac{p^M(\tau)}{q(p_a, \tau - t)}\right) \underline{\omega}'(t)\ \mathrm{d}t\right)$$

which simplifies to:

$$W_a(\tau) := \begin{cases} \dfrac{(1-p_a)\left(1-e^{-\lambda_\ell p_a \tau}\right)}{\alpha} & \text{if } \tau \leqslant T_b \\ \dfrac{p_b(\alpha p_a - 1)\left(\frac{p_a-1}{\alpha p_a - 1}\right)^{\frac{1}{p_a}} e^{-\lambda_\ell \tau}}{\alpha} - p_a p_b + p_a & \text{otherwise.} \end{cases}$$

The calculation for the representative b-worker is similar. For any $\tau < T_b$, no b-worker is tried, so the expected wage of the representative b-worker is 0. For $\tau \geqslant T_b$, it is:

$$\frac{1}{\beta} \int_{T_b}^{\tau} p_b \left(1 - \frac{p^M(\tau)}{q(p_b, \tau - t)}\right) \underline{\omega}'(t) \, \mathrm{d}t.$$

Hence,

$$W_b(\tau) := \begin{cases} 0 & \text{if } \tau \leqslant T_b \\ \dfrac{(\alpha p_a - 1)\left(\left(\frac{p_a-1}{\alpha p_a-1}\right)^{\frac{p_b}{p_a}} e^{-\lambda_\ell p_b \tau} - p_b \left(\frac{p_a-1}{\alpha p_a-1}\right)^{\frac{1}{p_a}} e^{-\lambda_\ell \tau} + p_b - 1\right)}{\beta} & \text{otherwise.} \end{cases}$$

At the start of the horizon, there exists a wage gap between groups because $W_a(0) > 0 = W_b(0)$. Moreover, the wage gap persists over the entire horizon and it does not disappear even in the long run, as the following lemma shows. This is because even as $\tau \to +\infty$, there exist a mass of b-workers who never get tried.

Proposition C.3 (Persistent wage gap under breakdowns)**.** *Suppose that $p_a(\alpha + \beta) > 1$. In the limit $p_b \uparrow p_a$,*

1. *for any $\tau \in [0, T_b)$, the wage gap $W_a(\tau) - W_b(\tau)$ is strictly increasing.*

2. *for any $\tau \in [T_b, \infty)$, the wage gap $W_a(\tau) - W_b(\tau)$ is first strictly increasing and then strictly decreasing. The limit $\lim_{\tau \to \infty} (W_a(\tau) - W_b(\tau))$ is strictly positive.*

Proof. We further assume that $\alpha > 1$ and $\alpha p_a < 1$, so $T_b \in (0, \infty)$. If $\alpha < 1$, then $T_b = 0$. If $\alpha p_a > 1$, then $T_b = \infty$. The proofs for both cases are similar, so are omitted.

For any $\tau \in [0, T_b)$, the wage gap $W_a(\tau) - W_b(\tau)$ is simply $W_a(\tau)$, which is strictly increasing in τ.

For any $\tau \in [T_b, \infty)$, the wage gap is increasing in τ if and only if

$$\frac{(\alpha+\beta)\left(\frac{1-p_a}{1-\alpha p_a}\right)^{\frac{1}{p_a}-1} e^{-\lambda_\ell(1-p_a)\tau}}{\alpha} > 1.$$

The LHS is decreasing in τ, so this inequality holds when τ is small enough. Since the LHS equals zero when $\tau \to \infty$ and the inequality holds when $\tau = T_b$, the wage gap is first strictly increasing and then strictly decreasing. In the limit of $\tau \to \infty$, the wage gap is strictly positive:

$$\lim_{\tau\to\infty} (W_a(\tau) - W_b(\tau)) = \frac{(1-p_a)(\alpha p_a + \beta p_a - 1)}{\beta} > 0.$$

■

Figure 3 illustrates the wage dynamics and the wage gap characterized in Proposition C.3.

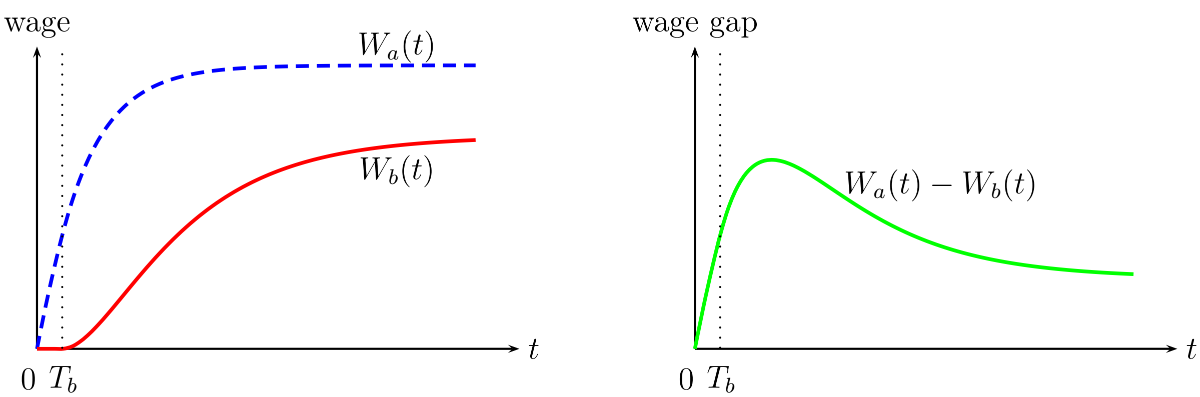

Figure 3: Wage dynamics and wage gap: $p_a = 1/2, \lambda_\ell = 1, \alpha = 5/4, \beta = 1$

If $p_a\alpha + p_b\beta \leqslant 1$ instead, all b-workers will obtain a chance in the long run. Even though for each $\tau \geqslant 0$ there exists a non-zero wage gap, in the limit the expected flow wages of the two groups converge.

D Proofs for Section 7.1

We consider here the case of a general $(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$. Based on whether the arrival of a signal is more likely to suggest a high or a low type, we distinguish two classes of learning environments:

(i) a signal is an *inconclusive breakthrough* if $\lambda_h > \lambda_\ell \geqslant 0$;

(ii) a signal is an *inconclusive breakdown* if $\lambda_\ell > \lambda_h \geqslant 0$.

The case of $\lambda_h = \lambda_\ell \geqslant 0$ corresponds to uninformative signals.

Proposition D.1 (Self-correcting property of inconclusive breakthroughs). *For any $\lambda_h > \lambda_\ell$, the two workers' payoffs converge as $p_a \downarrow p_b$.*

Proof. Let $U_i(p_a, p_b)$ be worker i's payoff given the belief pair (p_a, p_b). For any $p_a > p_b$, the principal first uses worker a for a period of length t^*. If no breakthrough occurs in $[0, t^*)$, the principal's belief toward worker a drops to p_b. Let $f(s)$ for $s \in [0, t^*)$ be the density of the random arrival time of the first breakthrough from worker a. We let $p_a(s)$ be the belief that $\theta_a = h$ if there is no breakthrough up to time s, and let $j(p_a(s))$ be the belief that $\theta_a = h$ right after the first breakthrough at time s. worker a's payoff is given by

$$\int_0^{t^*} f(s)\left(1 - e^{-rs} + e^{-rs} U_a(j(p_a(s)), p_b)\right) \mathrm{d}s + \left(1 - \int_0^{t^*} f(s)\,\mathrm{d}s\right)\left(1 - e^{-rt^*} + e^{-rt^*} U_a(p_b, p_b)\right).$$

worker b's payoff is given by

$$\int_0^{t^*} f(s) e^{-rs} U_b(j(p_a(s)), p_b)\,\mathrm{d}s + \left(1 - \int_0^{t^*} f(s)\,\mathrm{d}s\right) e^{-rt^*} U_b(p_b, p_b).$$

As $p_a \downarrow p_b$, t^* converges to zero. Both players' payoffs converge to $U_a(p_b, p_b) = U_b(p_b, p_b)$. ∎

Proposition D.2 (Spiraling property of inconclusive breakdowns). *For any $\lambda_h < \lambda_\ell$, the two workers' payoffs do not converge if $r^2 - (1 - 2p_a) r(\lambda_\ell - \lambda_h) - \lambda_h \lambda_\ell > 0$ or equivalently:*

$$\frac{\lambda_h}{\lambda_h + r} p_a + \frac{\lambda_\ell}{\lambda_\ell + r}(1 - p_a) < \frac{1}{2}.$$

Proof. Let $U_i(p_a, p_b)$ be worker i's payoff given the belief pair (p_a, p_b). We let $p_a(s)$ be the belief toward worker a if there is no breakdown up to time s, and let $j(p_a(s))$ be the belief toward him right after the first breakdown at time s.

For any $p_a > p_b$, the principal begins with worker a, and uses worker a exclusively if no breakdown occurs. We let $f(s) = p_a \lambda_h e^{-\lambda_h s} + (1 - p_a)\lambda_\ell e^{-\lambda_\ell s}$ be the density of the arrival time $s \in [0,\infty)$ of the first breakdown from worker a. Correspondingly, $1 - F(t)$ is the probability that no breakdown occurs before time t. For any $t > 0$, we can write worker a's payoff as follows:

$$\int_0^t f(s)\left\{1 - e^{-rs} + e^{-rs}U_a\left(j(p_a(s)), p_b\right)\right\} \mathrm{d}s + (1 - F(t))\left(1 - e^{-rt} + e^{-rt}U_a(p_a(t), p_b)\right).$$

We can write worker b's payoff as follows:

$$\int_0^t f(s)\left\{e^{-rs}U_b\left(j(p_a(s)), p_b\right)\right\} \mathrm{d}s + (1 - F(t))e^{-rt}U_a(p_a(t), p_b).$$

The payoff difference between a and b is:

$$\int_0^t f(s)\left(1 - e^{-rs} + e^{-rs}\left(U_a\left(j(p_a(s)), p_b\right) - U_b\left(j(p_a(s)), p_b\right)\right)\right) \mathrm{d}s \\ + (1 - F(t))\left(1 - e^{-rt} + e^{-rt}\left(U_a(p_a(t), p_b) - U_b(p_a(t), p_b)\right)\right).$$

We claim that $U_a(p_a, p_b) - U_b(p_a, p_b) \geqslant -1$ since $U_i(p_a, p_b)$ is in the range $[0, 1]$ for any i, p_a, p_b. Therefore, the payoff difference is at least:

$$H(t) := \int_0^t f(s)\left(1 - 2e^{-rs}\right) \mathrm{d}s + (1 - F(t))\left(1 - 2e^{-rt}\right).$$

Since we can choose any t, the payoff difference is at least $\sup_t H(t)$. Note that $H(0) = 0$, and $H(\infty) > 0$ if and only if $r^2 - (1 - 2p_a)r(\lambda_\ell - \lambda_h) - \lambda_h\lambda_\ell > 0$. ■

References

Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, 30(2): 175–187.

Ali, S. Nageeb, and Ce Liu. 2020. "Conventions and Coalitions in Repeated Games."

Altonji, Joseph G. 2005. "Employer Learning, Statistical Discrimination and Occupational Attainment." *American Economic Review: Papers and Proceedings*, 95(2): 112–117.

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

Antonovics, Kate, and Limor Golan. 2012. "Experimentation and job choice." *Journal of Labor Economics*, 30(2): 333–366.

Arrow, Kenneth. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets.* , ed. O. Ashenfelter and A. Rees, 3–33. Princeton University Press.

Baron, James N., and David M. Kreps. 1999. *Strategic Human Services: Frameworks for General Managers.* New York:Wiley.

Bergemann, Dirk, and Juuso Valimaki. 1996. "Learning and Strategic Pricing." *Econometrica*, 64(5): 1125–1149.

Bergemann, Dirk, and Ulrich Hege. 2005. "The Financing of Innovation: Learning and Stopping." *The RAND Journal of Economics*, 36(4): 719–752.

Bertrand, Marianne, and Esther Duflo. 2017. "Handbook of Field Experiments." , ed. Esther Duflo and Abhijit Banerjee Vol. 1, Chapter Field Experiments on Discrimination. Elsevier.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94(4): 991–1013.

Blank, Rebecca. 2005. "Tracing the Economic Impact of Cumulative Discrimination." *American Economic Review: Papers and Proceedings*, 95(2): 99–103.

Blank, Rebecca, Marilyn Dabady, and Connie Citro. 2004. "Measuring racial discrimination." National Research Council The National Academies Press, Washington DC.

Blume, Lawrence E. 2006. "The Dynamics of Statistical Discrimination." *The Economic Journal*, 116: F480–F498.

Bohren, J. Aislinn, Alex Imas, and Michael Rosenberg. 2019. "The Dynamics of Discrimination: Theory and Evidence." *American Economic Review*, 109(10): 3395–3436.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2019. "Inaccurate Statistical Discrimination."

Bonatti, Alessandro, and Johannes Hörner. 2011. "Collaborating." *American Economic Review*, 101(2): 632–63.

Bose, Gautam, and Kevin Lang. 2017. "Monitoring for Worker Quality." *Journal of Labor Economics*, 35(3): 755–785.

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

Cornell, Bradford, and Ivo Welch. 1996. "Culture, Information, and Screening Discrimination." *Journal of Political Economy*, 104(3): 542–571.

Deb, Rahul, Matthew Mitchell, and Mallesh M. Pai. 2019. "Our Distrust is Very Expensive." *working paper.*

Fang, Hanming, and Andrea Moro. 2011. "Chapter 5 - Theories of Statistical Discrimination and Affirmative Action: A Survey." In . Vol. 1 of *Handbook of Social Economics*, 133 – 200. North-Holland.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *The Quarterly Journal of Economics*, 111(4): 1007–1047.

Felli, Leonardo, and Christopher Harris. 1996. "Learning, Wage Dynamics, and Firm-Specific Human Capital." *Journal of Political Economy*, 104(4): 838–868.

Fershtman, Daniel, and Alessandro Pavan. 2020. ""Soft" Affirmative Action and Minority Recruitment."

Foster, Dean P., and Rakesh V. Vohra. 1992. "An Economic Argument for Affirmative Action." *Rationality and Society*, 4(2): 176–188.

Ghosh, Arpita, and Patrick Hummel. 2012. "Learning and Incentives in User-Generated Content: Multi-Armed Bandits with Endogenous Arms."

Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians." *American Economic Review*, 90(4): 715–741.

Gu, Jiadong, and Peter Norman. 2020. "A Search Model of Statistical Discrimination." *Available at SSRN 3575199*.

Guo, Yingni. 2016. "Dynamic Delegation of Experimentation." *American Economic Review*, 106(8): 1969–2008.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2016. "Optimal Contracts for Experimentation." *The Review of Economic Studies*, 83(3): 1040–1091.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2017. "Contests for Experimentation." *Journal of Political Economy*, 125(5): 1523–1569.

Hörner, Johannes, and Larry Samuelson. 2013. "Incentives for experimenting agents." *The RAND Journal of Economics*, 44(4): 632–663.

Jacobs, David. 1981. "Toward a Theory of Mobility and Behavior in Organizations: An Inquiry into the Consequences of Some Relationships Between Individual Performance and Organizational Success." *American Journal of Sociology*, 87(3): 684–707.

Keller, Godfrey, and Sven Rady. 2010. "Strategic experimentation with Poisson bandits." *Theoretical Economics*, 5: 275–311.

Keller, Godfrey, and Sven Rady. 2015. "Breakdowns." *Theoretical Economics*, 10(1): 175–202.

Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic Experimentation with Exponential Bandits." *Econometrica*, 73(1): 39–68.

Kim, Young-Chul, and Glenn C. Loury. 2018. "Collective Reputation and the Dynamics of Statistical Discrimination." *International Economic Review*, 59(1): 1–16.

Kurtulus, Fidan Ana. 2016. "The Impact of Affirmative Action on the Employment of Minorities and Women: A Longitudinal Analysis Using Three Decades of EEO-1 Filings." *Journal of Policy Analysis and Management*, 35(1): 34–66.

Lange, Fabian. 2007. "The Speed of Employer Learning." *Journal of Labor Economics*, 25(1): 1–35.

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

Lo Sasso, Anthony T., Michael R. Richards, Chiu-Fang Chou, and Susan E. Gerber. 2011. "The $16,819 Pay Gap For Newly Trained Physicians: The Unexplained Trend Of Men Earning More Than Women." *Health Affairs*, 30(2).

Mansour, Hani. 2012. "Does Employer Learning Vary by Occupation?" *Journal of Labor Economics*, 30(2): 415–444.

Miller, Amalia R., and Carmit Segal. 2012. "Does Temporary Affirmative Action Produce Persistent Effects? A Study of Black and Female Employment in Law Enforcement." *The Review of Economics and Statistics*, 94(4): 1107–1125.

Miller, Conrad. 2017. "The Persistent Effects of Temporary Affirmative Action." *American Economic Journal: Applied Economics*, 9(3): 152–190.

Moro, Andrea, and Peter Norman. 2004. "A general equilibrium model of statistical discrimination." *Journal of Economic Theory*, 114: 1–30.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology*, 108(5): 937–975.

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review*, 62(4): 659–661.

Sarsons, Heather. 2019. "Interpreting Signals in the Labor Market: Evidence from Medical Referrals."

Shapley, Lloyd S, and Martin Shubik. 1971. "The assignment game I: The core." *International Journal of game theory*, 1(1): 111–130.

Strulovici, Bruno. 2010. "Learning While Voting: Determinants of Collective Experimentation." *Econometrica*, 78(3): 933–971.