Early-Career Discrimination: Spiraling or Self-Correcting?

Arjada Bardhi Yingni Guo Bruno Strulovici*

August 3, 2026

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. In *breakdown learning environments* that primarily track on-the-job failures, such discrimination spirals into a substantial lifetime earnings gap for groups of comparable productivity, whereas in *breakthrough learning environments* that track successes, early-career discrimination can be self-corrected, so comparable groups obtain comparable lifetime earnings. This contrast persists in large labor markets and with flexible wages, inconclusive learning, and misspecified employer beliefs.

JEL: C78, D83, J71

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 shown to rely on observable characteristics—such as a worker's race or gender—as statistical proxies for the worker's productivity.[1] Such early-career statistical discrimination determines who makes the first cut when job opportunities are scarce. Even if different social groups have comparable productivity distributions, groups that make the first cut may be systematically prioritized.

Does the impact of such early-career discrimination on workers' earnings vanish or intensify over time? One plausible conjecture is that social groups of comparable productivity obtain comparable lifetime earnings: over time, employers learn about workers' productivity from observing their

*Bardhi: Department of Economics, New York University; email: arjada.bardhi@nyu.edu. Guo: Department of Economics, Duke University; email: yingni.guo@duke.edu. Strulovici: Department of Economics, Northwestern University; email: b-strulovici@northwestern.edu. We thank Attila Ambrus, Peter Arcidiacono, Heski Bar-Isaac, Aislinn Bohren, Yeon-Koo Che, Hanming Fang, Rachel Kranton, Kevin Lang, Erik Madsen, Peter Norman, Wojciech Olszewski, Alessandro Pavan, Giorgio Primicieri, Anja Prummer, Heather Sarsons, Curtis Taylor, Huseyin Yildirim, and seminar audiences at various seminars and workshops for valuable feedback.

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

performance and reallocate opportunities accordingly. However, an opposite conjecture suggests that comparable groups may fare drastically differently: when early opportunities to perform are pivotal to a worker's career progression, workers favored early on fare substantially better.

This paper shows that the way employers learn about workers' productivity makes a critical difference for which conjecture prevails. In environments that primarily track on-the-job successes, early-career discrimination has only minor consequences for workers' later job opportunities and lifetime earnings. In environments that primarily track on-the-job failures, in contrast, early-career discrimination significantly affects workers' lifetime prospects. Moreover, the adverse effect on workers who are discriminated against intensifies with job scarcity: the scarcer jobs are relative to the size of the workforce, the higher the inequality between groups. Our analysis thus suggests a classification of learning environments that predicts whether, and if so, under which circumstances the impact of early-career discrimination vanishes or gets amplified over time.

This contrast between learning environments persists even with flexible wages—a possibility that we formalize through a dynamic two-sided matching model. Even between highly comparable social groups, environments that track failures induce a substantial delay in employment and a significantly lower wage path for groups that are disfavored at the outset.

Model. We study labor markets in which (i) workers from different groups compete for 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 (2022) studies one such market in which male and female surgeons compete for referrals from physicians. Physicians learn about surgeons' abilities from surgeries they performed in the past. Sarsons (2022) documents comparable ability distributions for male and female surgeons: the average ability is only slightly lower for female surgeons in her sample.[3] We investigate the consequences of such a small difference.

We begin with a *small market* that features one employer and two workers identified by their respective social groups, a and b. A worker's productivity is either high or low, and worker a is ex ante more likely than worker b to have high productivity. In each instant, the employer allocates the task to one of the two workers—similar to a physician choosing a surgeon for referral—or takes an outside option if the expected productivity of both workers is too low. The employer's flow payoff increases in the productivity of the employed worker, whereas workers benefit from being allocated the task regardless of their productivity. In Section 3, we analyze a *large market* with a continuum of workers from each group and a continuum of employers.

The employer learns about a worker's productivity from their performance. We contrast two learning environments: *breakthrough* learning and *breakdown* learning. In the breakthrough environment, a high-productivity worker generates successes, which we call breakthroughs, at randomly distributed times, whereas a low-productivity worker generates nothing. In the breakdown

[2]These stylized features tractably capture a more general setting in which (i') some tasks are more desirable than others and desirable tasks are in limited supply, (ii') workers who perform desirable tasks reveal more about their productivity than do workers who are either employed in other tasks or unemployed, and (iii') groups do not necessarily have comparable productivity distributions. Our focus on groups with comparable productivity distributions highlights the role played by the learning environment in the dynamics of statistical discrimination.

[3]See Section 2.2.2 and Figure 1 in Sarsons (2022).

environment, a low-productivity worker generates failures, which we call breakdowns, at randomly distributed times, whereas a high-productivity worker generates nothing. We also analyze mixed learning environments that combine both breakthroughs from high-productivity workers and breakdowns from low-productivity workers; the relative frequency of the two determines whether the environment is *breakthrough-salient* or *breakdown-salient*.

The learning environment can be viewed 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). Star jobs—such as high-stakes salespeople, entertainers, and athletes—feature infrequent but large successes. By contrast, guardian jobs—routine surgeons, airline pilots, prison guards, among others—are marked by infrequent but costly failures.

Main results. In both learning environments, the employer first allocates the task to worker a, who has a higher expected productivity. 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 drops to that of worker b. From this point onward, the employer treats the two workers equally. The length of this grace period over which the task is allocated exclusively to worker a reflects the difference between the two workers' expected productivities at the start. The smaller this initial difference, the shorter the grace period for worker a, and the smaller the first-hire advantage of worker a. As this difference shrinks to zero, so does the advantage of worker a. The breakthrough environment is thus *self-correcting*.

In the breakdown environment, in contrast, the absence of a breakdown from worker a makes the employer more optimistic about the worker's productivity. Therefore, the employer allocates the task exclusively to worker a until a breakdown occurs. Worker b gets a chance to perform the task only if worker a has low productivity and misperforms on the task. As a result, worker b's expected lifetime earnings are only a fraction of worker a's. We show that even if worker a's productivity distribution is only slightly superior ex ante, this small initial difference spirals into a large payoff inequality. This *spiraling* effect persists *even as learning by the employer becomes arbitrarily fast*. It can explain why societies struggle to eliminate inequality in labor markets.

This contrast between breakthrough and breakdown environments extends to mixed learning environments with both breakthroughs and breakdowns. Any breakthrough-salient environment—in which breakthroughs arrive faster than breakdowns—continues to be self-correcting, whereas any breakdown-salient environment gives rise to spiraling. However, the contrast is now more nuanced: fixing the total arrival rate but varying the arrival rate of breakthroughs relative to breakdowns, the payoff gap between the two workers changes smoothly as the environment shifts continuously from pure breakthrough learning to pure breakdown learning. The more comparable the groups become, the more pronounced is the contrast between breakthrough-salient and breakdown-salient environments: the payoff gap converges to zero for any breakthrough-salient environment but converges to the strictly positive gap of a pure breakdown environment for any breakdown-salient environment.

We further explore this contrast in the large market. We show that the key determinant of the spiraling effect in the breakdown environment is the scarcity of tasks relative to the size of the workforce. As tasks become scarcer, the inequality between groups increases. Hence, while all groups suffer from a decrease in labor demand during economic downturns, *groups that are discriminated against will suffer disproportionately more.*

One might conjecture that flexible wages will eliminate inequality between groups of comparable productivity. For instance, Becker (1957) and later Flanagan (1978) have argued that with flexible wages, the wage differential should equalize the employment rates across groups. To evaluate this conjecture, we introduce flexible wages into the large market. From a methodological standpoint, we develop a dynamic two-sided matching model that incorporates both learning and flexible wages, and show that the essentially unique stable stage-game matching is *dynamically stable* (Ali and Liu, 2026).

We find that flexible wages do not resolve the severe differential treatment of comparable groups in the breakdown environment. Intuitively, they do not overcome the tension caused by task scarcity: when only a subset of workers can be hired, those who generate higher expected surplus get hired first. Hence, as in the case of fixed wages, b-workers are not given a chance to perform unless and until sufficiently many a-workers have experienced breakdowns. Such a delay further implies that employers learn more about a-workers than b-workers. Hence, employed a-workers (i.e., those who have not generated breakdowns) earn a higher wage than employed b-workers. Breakdown learning thus results in substantial gaps in wages and earnings between groups of almost identical expected productivity, where the wage gap reflects differences in pay rates conditional on employment, whereas the earnings gap reflects both the wage differential (i.e., pay per task) and the employment gap (i.e., access to tasks). Figure 1 illustrates the predicted paths of average wages and those of average flow earnings.[4] Both the gap in average wages and that in average flow earnings widen early on in the workers' careers and persist for a substantial amount of time. If tasks are sufficiently scarce relative to workers, *the gap in flow earnings persists throughout the workers' careers.*

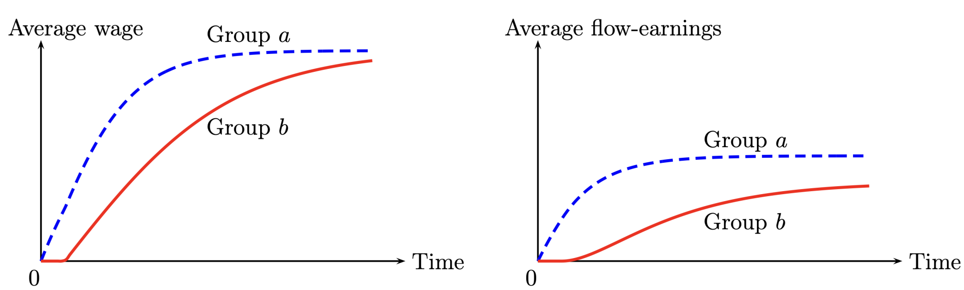


Figure 1: Average wage/flow earnings under breakdowns for groups of comparable productivity

[4]Average flow earnings are defined as the average payoff across both employed and unemployed workers. The average wage is taken across employed workers only, so it is higher than average flow earnings.

While our main characterization assumes a zero minimum wage, we show that the dynamics are similar with a negative minimum wage. Workers outbid each other to the point that the marginal-productivity worker among those hired in each instant is paid the negative minimum wage. Since the employment dynamics are identical to those under a zero minimum wage, b-workers face the same delay in employment.

Related literature

First and foremost, our paper contributes to the theoretical literature on statistical discrimination surveyed by Fang and Moro (2011) and Onuchic (2023). Phelps (1972) and the subsequent literature (e.g., Aigner and Cain (1977), Cornell and Welch (1996), and Fershtman and Pavan (2021)) assume a significant, exogenous difference between social groups, which gives rise to inequality between groups. In contrast, Arrow (1973) and the subsequent literature (e.g., Foster and Vohra (1992), Coate and Loury (1993), 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 within an equilibrium.[5]

Our approach differs from both of these strands of literature. First, we consider groups that share arbitrarily similar productivity. In the models building on Phelps (1972), inequality across groups disappears as the productivity difference vanishes, whereas our model shows how a vanishingly small difference can snowball into a large payoff gap. Second, in contrast to Arrow (1973), the across-group inequality that we uncover is not due to the existence of multiple equilibria. Third, most papers in both strands do not model group interaction, whereas in our model groups compete for tasks. From this standpoint our paper is related to Cornell and Welch (1996) in the first group and Moro and Norman (2004) in the second. However, these two papers consider static task allocation, whereas we explore the consequences of dynamic task allocation.

Our analysis also contributes to work on cumulative discrimination (Blank, Dabady and Citro (2004), Blank (2005)) and systemic discrimination (Bohren, Hull and Imas, 2025), which study how discrimination compounds over time, leading to persistent disparities as discrimination at one stage distorts the decisions and information available at the next. Our results predict that such forms of discrimination, even when driven by negligible initial direct discrimination, will be more severe in breakdown than in breakthrough environments. When systemic discrimination is taken to data, therefore, we predict heterogeneity by sector and type of job. Similar to Pallais (2014), our findings emphasize the informational value of entry-level jobs. In a two-stage tournament model, Drugov, Meyer and Möller (2024) show that organizations favor early strong performers, even when performance differences stem largely from luck. This bias gives early luck lasting impact. Both the bias and the persistence are amplified by agents' strategic effort.

Our results also speak to the employer learning literature (e.g., Farber and Gibbons (1996), Altonji and Pierret (2001)), interpreting the learning environment as an intrinsic feature of an

[5] Blume (2006) and Kim and Loury (2018) extend the static setup of Coate and Loury (1993) to incorporate generations of workers. In contrast, we examine a single generation of long-lived workers whose productivity is revealed gradually while performing tasks.

occupation. Related work by Altonji (2005), Lange (2007), Antonovics and Golan (2012), and Mansour (2012) models occupational differences in the frequency of signals, whereas we allow the direction of signals to vary across occupations and show its importance. Hurst, Rubinstein and Shimizu (2024) similarly find that discrimination varies with task characteristics; however, in our setting, tasks differ in learning environment, whereas in theirs, in social interaction. Relatedly, Bartoš et al. (2016) study attention discrimination in cherry-picking (e.g., labor) versus lemon-dropping (e.g., rental housing) markets. While their mechanism differs from ours, they similarly emphasize how the selection context and direction of learning shape allocation.

Our model leverages the tractability of Poisson bandits, which are used widely in strategic experimentation models (e.g., Keller, Rady and Cripps (2005), Keller and Rady (2010), Strulovici (2010), Keller and Rady (2015)).[6] Both Felli and Harris (1996) and this paper use the framework of multi-armed bandits to model labor market learning. Felli and Harris (1996) assume one worker and two employers, whereas our small market has one employer and two workers so job opportunities are in short supply.[7] We make two technical contributions to the literature on Poisson bandits. First, while most existing work focuses on the operator's payoff, we study the occupation measure of each arm—the analog of a worker's expected lifetime payoff. Second, we embed the Poisson learning model in a dynamic two-sided matching model with flexible wages, and we show that it admits a unique Markovian, dynamically stable matching. The model accommodates rich heterogeneity while remaining analytically transparent, enabling testable predictions about how the learning environment shapes wage dynamics and long-run inequality. The contrasting of breakthrough and breakdown learning adds to a recent literature that compares the implications of good-news learning and bad-news learning in various applications: Board and Meyer-ter-Vehn (2013) on reputational incentives, Halac and Prat (2016) on managerial attention, and Halac and Kremer (2020) on career concerns. Our mixed learning environment, which features both conclusive breakthroughs and conclusive breakdowns, is similar to that in Halac and Prat (2016), Che and Hörner (2018), and more recently Lizzeri, Shmaya and Yariv (2024).

This paper contributes to a growing literature that employs bandit models to study statistical discrimination. Li, Raymond and Bergman (2024) study an exploration-based screening algorithm that leads to higher diversity and quality of workers. Learning about minority groups can lag behind due to early negative signals from minority workers (Lepage, 2023), population imbalances (Komiyama and Noda, 2024), limited attention in search (Che, Kim and Zhong, 2019), and noisier evaluation of minority groups (Fershtman and Pavan, 2021). We show that learning about groups disfavored at the start can lag behind also due to the nature of employer learning.

[6] Other areas of applications include moral hazard (e.g., Bergemann and Hege (2005), Hörner and Samuelson (2013), Halac, Kartik and Liu (2016)), collaboration (e.g., Bonatti and Hörner (2011)), delegation (e.g., Guo (2016)), and contest design (e.g., Halac, Kartik and Liu (2017)).

[7] We explore workers' incentives to invest in productivity in Section 4.2. This section is related to Bergemann and Valimaki (1996), Felli and Harris (1996), and Deb, Mitchell and Pai (2022), since it also models bandit arms as strategic players. However, all these models assume that the quality of the arms is exogenously given, while it is endogenously determined in our Section 4.2.

2 Small market

2.1 Framework

Players and types. We consider a small labor market with one employer ("she") and two workers (each "he"). Time $t \in [0, \infty)$ is continuous, and the discount rate is $r > 0$. 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 $t = 0$, workers' types are drawn independently of each other. Worker i's type θ_i is either high ($\theta_i = h$) or low ($\theta_i = \ell$). The prior probability that worker i has a high type is $p_i \in (0, 1)$. Neither the employer nor the workers know the workers' realized types. They only share the belief (p_a, p_b), and all learning is public. In particular, the assumption that the worker does not know his type is consistent with our focus on early-career dynamics, in which it is plausible that workers do not know their types yet. Such an assumption is standard in models of learning in labor markets, such as Felli and Harris (1996) and Altonji and Pierret (2001). We interchangeably refer to p_i as the prior belief for worker i or as worker i's expected productivity at time 0. We assume that worker a is ex ante more productive: $p_a > p_b$. Our focus is on groups with comparable expected productivity, i.e., when 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 a known outside option, such as a worker with some known productivity or the value from reallocating resources to other organizational goals. A worker obtains a flow payoff $w > 0$ whenever he is assigned the task. Otherwise, his flow payoff is zero. We interpret w as the fixed wage for a worker who performs the task and normalize it without loss to $w = 1$. The employer obtains a flow payoff $v > 0$ if the task is allocated to a high-type worker, and a flow payoff normalized to zero if it is allocated to a low-type worker. These payoffs can be interpreted as the employer's net payoffs after the wage w is paid. If the employer allocates the task to the safe arm, she earns a flow payoff $s \in (0, v)$.[8]

Learning by allocating tasks. Learning about a worker's type proceeds via conclusive Poisson signals. If worker i is allocated the task over the interval $[t, t + \mathrm{d}t)$ and his type is $\theta_i = h$, a public *breakthrough* arrives with probability $\lambda_h \mathrm{d}t$ and no breakthrough arrives with the complementary probability. If $\theta_i = \ell$ instead, a public *breakdown* arrives with probability $\lambda_\ell \mathrm{d}t$ and no breakdown arrives with complementary probability. Thus, a *learning environment* is characterized by a pair of type-dependent arrival rates $(\lambda_h, \lambda_\ell)$. A breakthrough reveals a high type and a breakdown reveals a low type. We assume for simplicity that the employer's payoffs are observed only at the end of the horizon, so that the employer learns only from the Poisson signals.[9] Based on which

[8] We multiply players' lifetime payoffs by r, as in Keller, Rady and Cripps (2005). This normalizes the employer's lifetime payoff from always hiring a high type to v and a worker's lifetime payoff from always being hired to 1.

[9] This formulation is, however, equivalent to an alternative formulation in which (i) the employer learns through observable payoffs, (ii) type h generates a lump-sum payoff b_h at arrival rate λ_h and type ℓ generates a lump-sum payoff $b_\ell \neq b_h$ at arrival rate λ_ℓ. Any nonzero values b_h, b_ℓ would be consistent with our analysis as long as $\lambda_h b_h > s > \lambda_\ell b_\ell$. Our chosen formulation makes it easier to compare payoffs across the learning environments.

type is revealed more quickly, we distinguish between two classes of environments:

(i) *breakthrough-salient environments*: $\lambda_h > \lambda_\ell \geqslant 0$;

(ii) *breakdown-salient environments*: $\lambda_\ell > \lambda_h \geqslant 0$.

The relative size of the arrival rates determines whether the employer's belief drifts down or up in the absence of any signal. In a breakthrough-salient (resp., breakdown-salient) environment, the employer becomes more pessimistic (resp., optimistic) that the worker has a high type. In the symmetric environment with $\lambda_h = \lambda_\ell$, such an absence is uninformative of the worker's type.

Central to our analysis are the environments in which only one type of signals is possible. If $\lambda_h > 0$ and $\lambda_\ell = 0$, we call the environment a *pure breakthrough environment* or simply a *breakthrough environment*. If $\lambda_h = 0$ and $\lambda_\ell > 0$, we call it a *pure breakdown environment* or simply a *breakdown environment*. If both $\lambda_h > 0$ and $\lambda_\ell > 0$, we say that the environment is *mixed*.[10]

Let $\underline{p}$ denote the belief threshold below which the employer switches to the safe arm, which is derived in Lemma A.1 and given by:

$$\underline{p} := \frac{rs}{rv + \max\{\lambda_h, \lambda_\ell\}(v - s)}.$$

This threshold depends only on $\max\{\lambda_h, \lambda_\ell\}$, the higher of the two arrival rates. As expected, $\underline{p}$ is lower than the myopic threshold s/v due to the value of learning for future allocation decisions. Hereafter, we assume that $p_i > \underline{p}$ for $i \in \{a, b\}$, so the employer prefers to experiment with both workers before turning to the safe arm.

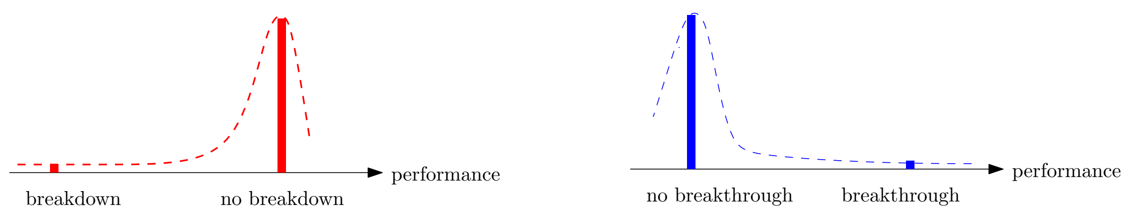
Discussion: A taxonomy of occupations

We interpret the learning environment as an intrinsic feature of how performance is monitored or evaluated at a given job. Following the job taxonomy suggested by Baron and Kreps (1999), breakthrough-salient environments are akin to star jobs and breakdown-salient environments are akin to guardian jobs. Star jobs are defined by infrequent but highly impactful successes from high-ability workers, where a single breakthrough can lead to outsized career rewards. High-stakes salespeople, professional athletes, startup founders and tech executives, and academic researchers in fields that reward singular discoveries are all examples of star jobs. Guardian jobs, on the other hand, are defined by the need to prevent rare but catastrophic failures, with career stability primarily dependent on avoiding mistakes. Examples of guardian jobs abound: airline pilots and air traffic controllers, compliance officers, firefighters and paramedics, surgeons in routine practices, nuclear plant operators, prison guards, structural engineers, and so on. Bose and Lang (2017) argue

[10]In mixed learning environments, signals are still conclusive. Section 4.3 shows that the main results remain qualitatively unchanged when allowing for inconclusive signals, i.e., when low-type (resp., high-type) workers also generate breakthroughs (resp., breakdowns) but at a lower rate than high-type (resp., low-type) workers. Despite some loss in tractability, we establish the self-correcting property in Proposition 4.1 and the spiraling property for sufficiently impatient players in Proposition 4.2.

that most nonmanagerial jobs are guardian jobs: reliability and mistake-free execution is more valuable than sporadic exceptional performance.

One way to empirically test whether a job is closer to a guardian or a star job is to look at the skewness of performance distributions. Baron and Kreps (1999) point out that the performance distributions differ substantially across star and guardian jobs: the distribution tends to be right-skewed for star jobs and left-skewed for guardian jobs. Figure 2 illustrates this. The dashed curves depict the probability densities of performance signals for a guardian job and a star job when the support of performance signals is an interval. The bars depict the probabilities when the performance signals are binary, as in our baseline model. "Breakdown" and "no breakdown" correspond to signals in a guardian job, whereas "breakthrough" and "no breakthrough" to those in a star job.[11]



(a) Guardian jobs (b) Star jobs

Figure 2: Distribution of performance (adapted from Figure 2-2 in Baron and Kreps (1999))

O'Boyle and Aguinis (2012) and Aguinis and Bradley (2015) show that in occupations centered around star performance, such as entertainment, athletics, and certain high-stakes business roles, the empirical distribution of performance is indeed right-skewed. This implies that "the majority of individuals are assumed to perform at an average level, with very few people actually achieving a level of performance that would place them in the category of being a star performer" (Aguinis and Bradley, 2015, p. 162). Similarly, empirical studies have consistently demonstrated that research performance in academia exhibits a right-skewed distribution, where a small fraction of scholars contribute disproportionately to the total research output (Bornmann (2024)).

By contrast, the most frequently used measures of surgical performance are composites of negative patient outcomes such as death, readmission, and post-operative complications (e.g., see large studies in Wallis et al. (2017) and Kelz et al. (2021) as well as Sarsons (2022)). These studies consistently show that the distribution of surgical performance appears to be tightly clustered around the lack of negative patient outcomes, with a minority of cases contributing to higher complication or mortality rates. Similarly, the most critical measures of pilot performance emphasize safety and procedural adherence, with incident rates, go-around rates, or FOQA rates being key performance indicators. This conceptual taxonomy of occupations deserves more empirical attention. The most direct way of testing the nature of an occupation would be to inspect the skewness of the performance distribution as in Figure 2.

[11]The bars do not condition on a worker's type, but they would look similar if the probabilities were conditional on a low type under breakdowns (Figure 2a) and conditional on a high type under breakthroughs (Figure 2b).

When is Poisson learning a good model of worker evaluation? Poisson learning captures environments with lumpy monitoring—where evaluation occurs irregularly, often triggered by high-visibility projects, infrequent inspections, or unexpected crises. Such monitoring leads to discrete jumps in the employer's belief. In many jobs, including those where performance is hard to observe or evaluate continuously, this kind of learning is the norm. But some occupations feature continuous monitoring, where employers receive a steady stream of performance data and update beliefs gradually. Examples include call center agents, data entry clerks, retail cashiers, and Uber drivers, whose work is tracked in real time through metrics like speed, accuracy, or customer satisfaction. Such jobs are less well captured by Poisson learning—a better way to model such incremental learning would be through Brownian signals. Still, such jobs are limited to environments where output is both highly standardized and closely measured. In most occupations—including many routine occupations like safety—monitoring remains intermittent and lumpy, making Poisson learning a broadly realistic model.

2.2 Contrast between the pure learning environments

We first observe a stark contrast between the two pure environments, and then generalize this contrast to mixed environments in Section 2.3. We compare the two workers' payoffs when they share arbitrarily similar expected productivities, and analyze how this comparison depends on whether the learning environment is pure breakthrough or pure breakdown.

Pure breakthrough learning. In each instant, the employer allocates the task to the worker with the higher expected productivity.[12] Since $p_a > p_b$, the employer first allocates the task to worker a. Because the belief that worker a has a high type drifts down as long as no breakthrough arrives, worker a is effectively given a grace period $[0, t^*)$ over which he is employed regardless of his performance, where 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 to him alone thereafter. Otherwise, starting from t^*, the employer splits the task equally between the two workers until either the belief drops down to $\underline{p}$ or a breakthrough occurs, so the workers obtain the same continuation payoff starting from t^*.

The hiring dynamics therefore go through two phases: a first phase during which worker a is hired exclusively, and a second phase during which the two workers are treated symmetrically starting from the symmetric belief p_b. Importantly, as p_b gets close to p_a, the first phase $[0, t^*)$ shrinks to zero. The probability that worker a generates a breakthrough before t^* converges to zero as well. Hence, as $p_b \uparrow p_a$, worker a's advantage vanishes and the two workers obtain similar expected payoffs. Therefore, we observe a *self-correcting property* of pure breakthrough learning:

[12]This is true in any learning environment that we consider because (i) workers' types are binary, and (ii) the arrival rates of signals and the employer's type-contingent flow payoffs are the same for both workers. For a fixed environment, workers' Gittins indices at each time t—which determine which worker is optimally chosen at time t—are given by the same increasing function of the workers' expected productivities $p_a(t), p_b(t)$ at time t (Gittins, Glazebrook and Weber, 2011).

a small difference in prior beliefs can result in only a small payoff advantage for worker a. This observation and the next one are generalized in Section 2.3 and proved formally in Appendix B.

Observation 2.1 (Self-correcting property of pure breakthrough learning)**.** *Let $\lambda_h > 0$ and $\lambda_\ell = 0$. As $p_b \uparrow p_a$, the expected payoff of worker b converges to that of worker a.*

Pure breakdown learning. Again, the employer first allocates the task to worker a. As long as worker a generates no breakdown, the employer becomes more optimistic that his type is high, so she continues to hire him. Once a breakdown is realized, the employer turns to worker b. If worker b also generates a breakdown, the employer resorts to the safe arm thereafter.

The lifetime payoff of a high-type worker a is one, since he is never fired, whereas that of a low-type worker a is $r/(\lambda_\ell + r)$. Once hired, the type-by-type continuation payoff of worker b is the same as that of worker a. However, worker b faces a delay in getting hired. He obtains an opportunity only if worker a is a low type—that is, with probability $(1 - p_a)$—and even then, it takes time for worker a's low type to be revealed, which further discounts worker b's payoff by a factor of $\lambda_\ell/(\lambda_\ell + r)$. Crucially, the delay faced by worker b is independent of how close p_b is to p_a: even if p_b is just slightly less than p_a, worker b obtains a substantially lower payoff than worker a. 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 unless and until he generates a breakdown. This stands in contrast to the pure breakthrough environment, in which worker a loses his preferential status if he fails to generate a breakthrough within the grace period.

Observation 2.2 (Spiraling property of pure breakdown learning)**.** *Let $\lambda_\ell > 0$ and $\lambda_h = 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a approaches*

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

Since groups a and b have comparable productivity distributions, even if the employer were blind to workers' group belonging and treated them identically, her payoff would be only slightly lower than what she attains when she observes group belonging. In the limit as $p_b \uparrow p_a$, her payoff would be the same with and without observing workers' group belonging. Therefore, making it more difficult for the employer to observe group belonging would equalize workers' payoffs without making the employer worse off.[13]

Pre-hiring signals offer a more limited remedy. When the employer observes a coarse signal about each worker before hiring—such as a pass-fail interview—worker b can be hired first only by outperforming worker a. Even as the signal becomes nearly perfectly accurate, this probability

[13] In the small market, the allocation of tasks is efficient but it leads to unequal payoffs between comparable groups in breakdown environments. Norman (2003) also identifies an equity-efficiency tradeoff, but through a different mechanism: group specialization helps employers form more informative priors and find matches more efficiently. While his setting makes inequality harder to fix without significant efficiency loss, ours allows simple remedies—like gender-blind hiring—when productivity gaps are small. For example, in a study of group-blind hiring practices, Goldin and Rouse (2000) show that blind orchestra auditions substantially increased the likelihood that female musicians advanced to the final round.

remains at most $p_a(1-p_a) \leq 1/4$ as $p_b \uparrow p_a$, so the initial disadvantage persists. With a sufficiently fine signal, such as a continuous Gaussian score, the probability that b is hired first can converge to 1/2, fully correcting spiraling. Because real-world pre-hiring signals are often coarse, however, spiraling persists even when they are quite accurate.

2.3 Contrast across mixed learning environments

We now establish that the sharp contrast between the two pure environments observed in Section 2.2 extends to mixed learning environments.

Self-correcting property of breakthrough-salient learning. Given that $\lambda_h > \lambda_\ell$, the employer's belief that worker a has a high type drifts down as long as no signal arrives. Similarly to the pure breakthrough environment, worker a is hired first and he is given a grace period $[0, t^*)$ over which to perform as long as he generates no breakdown. Here, t^* is the time it takes for the belief about worker a's type to drift down from p_a to p_b in the absence of a signal:

$$t^* = \frac{1}{\lambda_h - \lambda_\ell} \log \frac{p_a(1-p_b)}{(1-p_a)p_b} \tag{2}$$

If worker a generates a breakthrough before t^*, the employer allocates the task to him alone thereafter. If worker a generates a breakdown before t^*, the employer switches immediately to worker b. Otherwise, starting from t^*, the employer treats the two workers symmetrically, so they obtain the same continuation payoff starting from t^*. As in the pure breakthrough case, as $p_b \uparrow p_a$, the grace period $[0, t^*)$ shrinks to zero and so does worker a's advantage.

Proposition 2.1 (Self-correcting property of breakthrough-salient learning). *Let $\lambda_h > \lambda_\ell \geqslant 0$. As $p_b \uparrow p_a$, the expected payoff of worker b converges to that of worker a.*

Spiraling property of breakdown-salient learning. Given that $\lambda_\ell > \lambda_h$, in the absence of a signal the employer becomes more optimistic that the worker's type is high. She first allocates the task to worker a. If a breakthrough is generated, worker a is hired forever. If a breakdown is realized, the employer switches to worker b. In the absence of either signal, the employer continues to hire worker a. As in the pure breakdown environment, a high-type worker a is never fired, whereas a low-type worker a enjoys a substantial period of employment before being eventually fired. Because the reallocation of tasks is driven entirely by the occurrence of a breakdown, each player's expected payoff depends on λ_ℓ but not λ_h. Therefore, the resulting payoff ratio is the same as in the pure breakdown environment with the same λ_ℓ.

Proposition 2.2 (Spiraling property of breakdown-salient learning). *Let $\lambda_\ell > \lambda_h \geqslant 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a approaches*

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

The payoff ratio (3) has two components: (i) the factor $(1-p_a)$ reflects the fact that worker b obtains a chance only if worker a has a low type, and (ii) the factor $\lambda_\ell/(\lambda_\ell+r)$ reflects the expected time it takes for worker a's low type to be revealed. Even as the revelation of low types becomes instantaneous—that is, as $\lambda_\ell \to \infty$—the payoff ratio approaches $(1-p_a)$ rather than one due to a strong rank effect. Being the second hire is detrimental to worker b even under instantaneous employer learning, since worker b never obtains a chance if worker a has a high type.

Spiraling in the symmetric environment. In the environment with $\lambda_h = \lambda_\ell$, the employer's belief stays constant in the absence of a signal. The employer starts with worker a, and fires him if and only if a breakdown arrives revealing his low type. Therefore, the task allocation dynamics and the workers' payoffs are the same as in any breakdown-salient environment with the same λ_ℓ. The payoff ratio as $p_b \uparrow p_a$ is that in (3), so spiraling arises.

Remark 1 (Continuum of types). The contrast across learning environments goes beyond binary types. For example, suppose that worker i's type θ_i is uniform on $[\underline{\theta}_i, \underline{\theta}_i + \Delta]$, where $\underline{\theta}_a > \underline{\theta}_b > 0$ and $\Delta > 0$ parameterizes the uncertainty about the workers' types. As $\underline{\theta}_b \uparrow \underline{\theta}_a$, the expected productivity of worker b converges to that of worker a. If worker i is employed, a signal that reveals his type arrives according to a Poisson process, and its arrival rate is $\lambda(\theta_i)$. If $\lambda(\theta_i)$ increases in θ_i—e.g., $\lambda(\theta_i) = \lambda\theta_i$ for some $\lambda > 0$—the learning environment is a breakthrough-salient one and the employer becomes more pessimistic in the absence of a signal. If $\lambda(\theta_i)$ decreases in θ_i, the environment is a breakdown-salient one. The self-correcting property of breakthrough-salient learning and the spiraling property of breakdown-salient learning continue to hold. In particular, the upward belief drift in the absence of a signal gives worker a a payoff advantage that does not shrink to zero even as $\underline{\theta}_b \uparrow \underline{\theta}_a$.

2.4 Comparative statics with respect to the learning environment

Our analysis so far fixes the learning environment—that is, it fixes the arrival rates $(\lambda_h, \lambda_\ell)$—and examines the limit payoff ratio as $p_b \uparrow p_a$. This section performs the complementary exercise of fixing (p_a, p_b) with $p_a > p_b$ and examining how the payoff gap (i.e., the difference in the workers' expected payoffs) varies with the learning environment. We parametrize a family of learning environments by the total arrival rate $\lambda_h + \lambda_\ell = \lambda > 0$. This parametrization ranges from the pure breakdown environment $(\lambda_h, \lambda_\ell) = (0, \lambda)$ to the symmetric one $(\lambda_h, \lambda_\ell) = (\lambda/2, \lambda/2)$ and the pure breakthrough environment $(\lambda_h, \lambda_\ell) = (\lambda, 0)$. We let U_a and U_b denote worker a's and worker b's expected payoffs, respectively.

Taken together, Propositions 2.1 and 2.2 imply that the limit of the payoff gap as the groups become arbitrarily similar is discontinuous at the frontier between the breakdown-salient and breakthrough-salient environments: for $\lambda_h < \lambda/2$, the limit payoff gap as $p_b \uparrow p_a$ is strictly positive, whereas for $\lambda_h > \lambda/2$ the limit payoff gap is zero. Thus, an environment that is slightly breakdown-salient generates much higher inequality between arbitrarily similar groups than one that is slightly breakthrough-salient. We next show that in fact, for any given $p_a > p_b > \underline{p}$, the payoff gap is continuously differentiable in the arrival rates.

Lemma 2.1. *Fix $p_a > p_b > \underline{p}$ and $\lambda > 0$, and consider the family of environments*

$$\left\{(\lambda_h, \lambda_\ell) = \left(\frac{\lambda}{2} + \delta, \frac{\lambda}{2} - \delta\right) : \delta \in \left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right\}.$$

The payoff gap $(U_a - U_b)$ is continuously differentiable in δ.

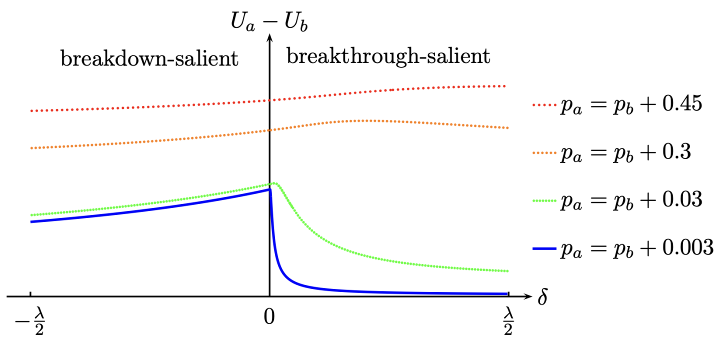


Figure 3: Payoff gap for four different levels of the prior difference. Parameter values are $\lambda = 2$, $s = 1/10$, $v = 1$, $r = 1$, and $p_b = 1/3$.

Figure 3 illustrates the payoff gap as a function of δ for four different levels of the prior difference $p_a - p_b$. As the prior difference gets smaller, the contrast between breakthrough-salient environments and breakdown-salient ones becomes more pronounced as well. The figure also shows that the payoff gap is much more sensitive to the prior difference in breakthrough-salient environments, dropping from significant levels down to zero as $(p_a - p_b) \downarrow 0$, than in breakdown-salient ones.

Within the class of breakdown-salient environments, the payoff gap increases in δ or, equivalently, decreases in λ_ℓ. We have argued in Section 2.3 that the workers' payoffs depend only on λ_ℓ in a breakdown-salient environment. The continuation payoff of each hired worker decreases in λ_ℓ, because low types are revealed faster. This shrinks the payoff gap between the two workers. Moreover, worker b expects to be hired earlier because a low-type worker a is revealed faster. This further shrinks the payoff gap. Altogether, the payoff gap decreases in λ_ℓ. Intuitively, faster learning about low types in breakdown-salient environments allows to weed out low-productivity workers from group a faster and reduce their initial advantage.

Within the class of breakthrough-salient environments, the payoff gap continues to increase in δ or, equivalently, decreases in λ_ℓ for environments sufficiently close to the symmetric one. The intuition can be understood by considering the environment that is only slightly breakthrough-salient. The employer's belief about worker a drifts down very slowly in the absence of any signal, so p_b is almost never reached. Thus, if worker a is fired, this is most likely due to a breakdown rather than as a result of worker a's expected productivity drifting down to the firing threshold.

The dynamics are thus similar to those of the breakdown-salient environment and, as in the breakdown-salient environment, the payoff gap increases in δ for a small enough δ. Proposition 2.3 formalizes these two observations.

Proposition 2.3. *Fix $p_a > p_b > \underline{p}$ and $\lambda > 0$, and consider the family of environments*

$$\left\{(\lambda_h, \lambda_\ell) = \left(\frac{\lambda}{2} + \delta, \frac{\lambda}{2} - \delta\right) : \delta \in \left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right\}.$$

(i) The payoff gap $(U_a - U_b)$ strictly increases in δ within the breakdown-salient environments.

(ii) There exists a maximal $\tilde{\delta} \in (0, \lambda/2]$ such that the payoff gap strictly increases in δ for $\delta \in (0, \tilde{\delta})$ within the breakthrough-salient environments.

Figure 3 shows that for a sufficiently large prior difference, it can be that $\tilde{\delta} = \lambda/2$, so the payoff gap strictly increases in δ throughout its domain. The figure moreover suggests that for $\delta > \tilde{\delta}$ the payoff gap strictly decreases in δ. Extensive numerical computations confirm such monotonicity. Two counteracting forces are at play as δ increases. On the one hand, if we consider the reallocation of tasks that is only driven by breakdowns, a lower λ_ℓ makes the payoff gap larger. However, increasing $(\lambda_h - \lambda_\ell)$ also shortens the grace period for worker a, which in return shrinks his payoff advantage. Our numerical computations show that if the second force dominates for some δ, then it dominates for any larger $\delta' > \delta$ as well.

3 Large market

The small market of the previous section features one employer choosing between two workers. We now focus on a large market with a continuum of workers from each group and a continuum of employers. Employers and workers are matched dynamically, tasks remain scarce, and all learning is public. The contrast between the breakthrough and the breakdown environments generalizes to this large market, regardless of whether wages are fixed or flexible. Notably, flexible wages are insufficient to prevent spiraling in the breakdown environment. Furthermore, in Appendix C.2.4, we show that this holds true even if the minimum wage is strictly negative, allowing workers to pay employers for early opportunities.

We introduce the framework in Section 3.1. Section 3.2 studies the dynamics of task allocation under fixed wages (i.e., wages are constant across time and across all matched employer-worker pairs), whereas Section 3.3 examines the case of flexible wages.

3.1 Framework

There is a unit mass of employers, a mass of size $\alpha > 1$ of workers from group a, and a mass of size $\beta > 0$ of workers from group b. All employers and workers are long-lived and 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. Crucially, there are more workers than tasks.[14] Employers are homogeneous. At $t = 0$, each worker's type is drawn independently from other workers' types. A worker from group $i \in \{a, b\}$ draws a high type with probability p_i. Additionally, there is a unit mass of identical safe arms that employers can take.

Stage-game matchings and stability. In this large market, a stage-game matching specifies: (i) how workers are matched to employers; and (ii) a wage for each matched pair. We use $k \in [0, \alpha+\beta]$ to index a worker and $j \in [0, 1]$ to index an employer. In the stage game, let $D_{kj} \in \{1, 0\}$ indicate whether worker k and employer j are matched to each other. If $D_{kj} = 1$, let W_{kj} denote the wage paid by employer j to worker k. Worker k's payoff is W_{kj} and employer j's expected payoff is $p_k v - W_{kj}$, where p_k is the probability that $\theta_k = h$.[15] All signals are public, so all employers share the same belief about each worker. If $D_{kj} = 0$ for all j, worker k is unmatched and gets zero payoff. If $D_{kj} = 0$ for all k, employer j takes a safe arm and gets a payoff of $s > 0$. Let $\mathcal{D}$ be the set of all such stage-game matchings. In the case of fixed wages, the wage for any matched pair is fixed at $w = 1$, so we only need to specify how workers are matched to employers. In the case of flexible wages, by contrast, any nonnegative wage is allowed.

We now apply the stability concept in Shapley and Shubik (1971) to our stage game. Given a stage-game matching (D, W), the pair (k, j) is called a *blocking pair* if each strictly prefers to be matched to the other at some allowable wage rather than follow (D, W).

Definition 1. A stage-game matching (D, W) is *stable* if (i) there is no matched employer j who strictly prefers to employ a safe arm instead of their matched worker, (ii) there is no matched worker k who strictly prefers to be unmatched, and (iii) there exists no blocking pair.

In Appendices C.1.1 and C.2.1, we show that each wage regime has an essentially unique stable stage-game matching: one under fixed wages, the other under flexible wages. In each case, the workers who are most productive and generate more surplus than the safe arm are matched.

Dynamic stability. In the dynamic game, a time-t history consists of all past matchings and realized signals until t. Let $\mathcal{H}_t$ be the set of all time-t histories, and $\mathcal{H} := \cup_{t \geqslant 0} \mathcal{H}_t$ the set of all histories. A *dynamic matching* μ specifies a lottery over stage-game matchings for any history, i.e., $\mu : \mathcal{H} \to \Delta(\mathcal{D})$. We define dynamic stability of a matching μ based on the solution concept of a perfect coalitional equilibrium in Ali and Liu (2026).[16] Given a dynamic matching μ, let $\mu(h_t)$ and μ_{h_t} respectively denote the stage-game matching and the continuation matching at h_t.

Definition 2. A dynamic matching μ is *dynamically stable* if at every t and every history $h_t \in \mathcal{H}_t$,

(i) there exists no matched employer j under $\mu(h_t)$ who strictly prefers to take a safe arm over the time window $[t, t + \mathrm{d}t)$, for some $\mathrm{d}t > 0$, and then revert to $\mu_{h_{t+\mathrm{d}t}}$;

[14] The assumption that $\alpha > 1$ simplifies the exposition since only workers from group a will be matched at $t = 0$. However, our results hold qualitatively for $\alpha < 1$ as well, provided that tasks are scarce, i.e., $\alpha + \beta > 1$.

[15] In the small market, v and 0 are the employer's payoffs from hiring, respectively, a high type and a low type *after* the fixed wage is paid. In the large market, we use them to represent an employer's payoffs *before* any wage is paid. This notation simplifies exposition within each market.

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

(ii) there exists no matched worker k under $\mu(h_t)$ who strictly prefers to be unmatched over $[t, t + \mathrm{d}t)$, for some $\mathrm{d}t > 0$, and then revert to $\mu_{h_{t+\mathrm{d}t}}$;

(iii) there exist no worker-employer pair (k, j) who strictly prefer to be matched to each other at some allowable wage over $[t, t + \mathrm{d}t)$, for some $\mathrm{d}t > 0$, and then revert to $\mu_{h_{t+\mathrm{d}t}}$.

In Appendices C.1.2 and C.2.2, we show that prescribing the essentially unique stable stage-game matching after each history is dynamically stable under both fixed and flexible wages. Let μ^* denote this dynamic matching. We focus on μ^* for three appealing properties. First, μ^* is dynamically stable for any discount rate $r > 0$ and remains dynamically stable even in the face of heterogeneity in players' discount rates. Second, μ^* is Markovian, in the sense that the matching at history h_t depends only on the distribution of workers' expected productivities at h_t. Third, Proposition C.1 in Appendix C.2.2 establishes that μ^* is the unique Markovian, dynamically stable matching under flexible wages. Sections 3.2 and 3.3 explore the properties of μ^* under fixed and flexible wages, respectively.[17]

3.2 Fixed wages

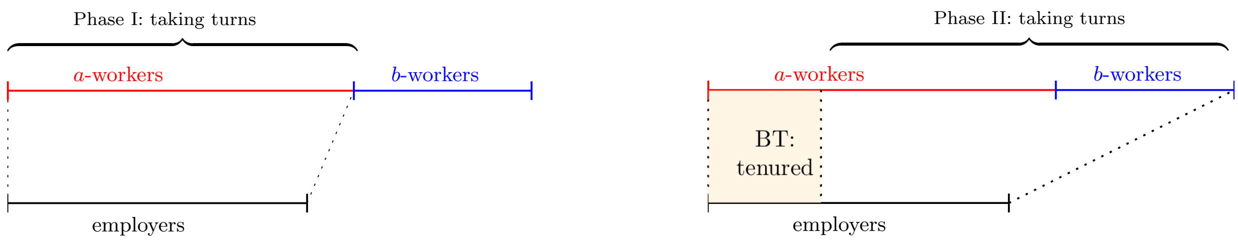
The intuition from the small-market analysis in Section 2 extends to the large market with fixed wages: comparable groups still earn similar payoffs in the breakthrough environment but diverge sharply in the breakdown one. Beyond extending the earlier results, the large-market analysis highlights how task scarcity amplifies payoff gaps—as tasks become scarcer relative to workers, inequality in the breakdown environment grows.

Broad hiring under breakthrough learning. As in Section 2.2, task allocation under breakthrough learning unfolds in two phases (Figure 4). In Phase I, all a-workers take turns to perform tasks, initially each getting a share $1/\alpha$ of a task. Those who generate a breakthrough "secure their job" and are allocated tasks at all future times. For other a-workers, employers' belief declines until reaching p_b, at which point a- and b-workers are viewed as equally productive. In Phase II, all remaining a-workers and all b-workers take turns to perform the tasks that have not been assigned to workers who have already generated a breakthrough.

Breakthrough learning thus encourages broad sampling of workers, echoing 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. (Baron and Kreps (1999), p. 28-29)

[17] While μ^* efficiently allocates tasks among workers of groups a and b, it does contain some inefficiency in that it switches to the safe arm too early relative to the utilitarian optimum. A social planner would continue employing a worker even when their myopic expected surplus falls below that of the safe arm in order to acquire information. Because decentralized firms do not internalize this informational externality, μ^* features under-experimentation relative to the social optimum.



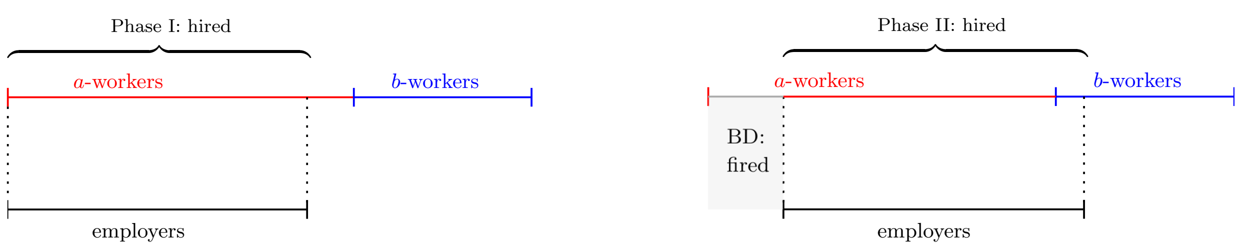
(a) Phase I: All a-workers hired (b) Phase II: All workers hired

Figure 4: Broad hiring under breakthrough learning

This broad-hiring dynamic promotes equity: as $p_b \uparrow p_a$, employers quickly extend their search to group b, so the payoffs of group b converge to those of group a. Hence, the self-correcting property extends to the large market.

Proposition 3.1. *Let $\lambda_h > \lambda_\ell = 0$, $\alpha > 1$, and $\beta > 0$. In the dynamic matching μ^*, the expected payoff of an a-worker converges to that of a b-worker as $p_b \uparrow p_a$.*

Narrow hiring under breakdown learning. Breakdown learning, in contrast, leads to sluggishness in experimenting with new workers: once hired, a worker remains employed until generating a breakdown. This inertia disproportionately harms group b regardless of how close p_b is to p_a, thus extending the insight of Observation 2.2 to large markets.



(a) Phase I: Unit mass of a-workers hired (b) Phase II: b-workers hired with delay

Figure 5: Narrow hiring under breakdown learning

Figure 5 illustrates the two phases of task allocation. Initially, each employer hires an a-worker, who remains employed unless a breakdown occurs. If so, he is replaced by another a-worker—until none remain. So b-workers must wait for their turn until all a-workers have been tried *and* sufficiently many of them have generated breakdowns. This delay persists even as $p_b \uparrow p_a$, keeping b-workers' payoffs strictly below those of a-workers.

Proposition 3.2. *Let $\lambda_\ell > \lambda_h = 0$, $\alpha > 1$, $\beta > 0$. In the dynamic matching μ^*, 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.*

Task scarcity magnifies inequality under breakdown learning. In this large market, α and β parametrize not only group sizes but also the relative scarcity of the unit mass of tasks. By varying α and β, we explore how inequality among groups varies with relative task scarcity.

Proposition 3.3 shows that under breakdown learning, inequality worsens as either α or β increases, holding task supply fixed.

Proposition 3.3 (Inequality increases in task scarcity under breakdown learning)**.** *Let $\lambda_\ell > \lambda_h = 0$, $\alpha > 1$, and $\beta > 0$. In the dynamic matching μ^*, as $p_b \uparrow p_a$, the limiting ratio of the expected payoff of an a-worker to that of a b-worker increases in both α and β.*

Increasing β intensifies competition within group b but does not affect the payoff of a-workers. By contrast, increasing α hurts both groups: it intensifies competition within group a and delays access to group b. Since each added a-worker uniformly delays all b-workers, the effect is harsher on group b. Thus, the scarcer tasks are relative to either group's labor supply, the greater the inequality between groups.

3.3 Flexible wages

We now incorporate flexible wages into the large market. We first explain how matches and wages are determined in a stable stage-game matching, and why prescribing this matching after each history is dynamically stable. We then show that flexible wages do not eliminate spiraling in the breakdown environment. To measure its severity, we derive closed-form expressions for key outcomes such as the employment rate, average wage, and average flow earnings by group. Lastly, we argue that relaxing limited liability still fails to prevent spiraling.

Stable stage-game matching. In Appendix C.2.1, we characterize the stable stage-game matching for any distribution G of workers' expected productivities. The allocation of workers to employers is statically efficient. For a given G, let $p^*(G)$ denote the least productive worker among the unit mass of the most productive workers. The marginal productivity is given by $p^M(G) := \max\{p^*(G), s/v\}$, determining which workers are matched: those with expected productivity above $p^M(G)$ are matched, while those below are not. Wages take a strikingly simple form—each matched worker with expected productivity p earns $(p - p^M(G))v$, reflecting the value they add relative to a marginal-productivity worker. Meanwhile, marginal-productivity worker(s) and unmatched workers earn zero. Notably, all employers earn the same profit of $p^M(G)v$, which is at least s. The left panel of Figure 6 illustrates workers' earnings as a function of p for a given p^M.

Why is μ^* dynamically stable? In a dynamic setting, the distribution G of workers' expected productivities evolves endogenously as their types are learned, resulting in a stochastic game with G as the state variable. Without learning, G would remain constant, and repeating the stable stage-game matching would be trivially dynamically stable. However, learning creates ripple effects in the evolution of the productivity distribution, rendering the following result far less straightforward.

Proposition 3.4. *In both the pure breakthrough environment and the pure breakdown one, μ^* is dynamically stable.*

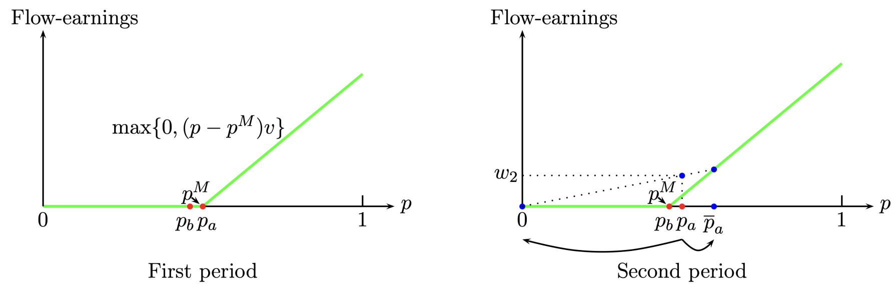


Figure 6: A two-period example with $\alpha = \beta = 1$

The proof verifies conditions (i)-(iii) in Definition 2; we elaborate here on the intuition behind each condition holding. Let $G(h_t)$ denote the distribution of expected productivities at history h_t. If employer j is matched under $\mu^*(h_t)$, deviating to a safe arm over $[t, t + \mathrm{d}t)$ results in weakly lower flow profits during that period, as stable stage-game matchings guarantee each employer at least s. Moreover, j's continuation payoff at $h_{t+\mathrm{d}t}$ remains unchanged, as each player in a large continuum market has a negligible impact on the state $G(h_{t+\mathrm{d}t})$. This observation mirrors general equilibrium logic: each player optimizes, taking the evolution of G as given, while the collective actions of all players determine its evolution.

If worker k is matched under $\mu^*(h_t)$, deviating to being unmatched over $[t, t + \mathrm{d}t)$ results in weakly lower flow earnings during that period, as stable stage-game matchings guarantee each worker at least zero earnings. While this deviation does not affect the state $G(h_{t+\mathrm{d}t})$, it reduces learning about worker k's type. Slower learning can only reduce k's continuation payoff, as a worker's flow earnings at any moment are a convex function of their expected productivity.

Lastly, we argue that no worker-employer pair (k, j) strictly prefers to match at some nonnegative wage over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$. We consider two cases based on whether k is already employed under $\mu^*(h_t)$:

(i) If k is employed, blocking with another employer j does not affect k's continuation payoff at $(t + \mathrm{d}t)$ since learning about k's type occurs identically regardless of who employs him. Moreover, since all employers earn the same flow profit over $[t, t + \mathrm{d}t)$ under $\mu^*_{h_t}$, no wage can yield both j and k strictly higher flow payoffs over $[t, t + \mathrm{d}t)$ than not blocking.

(ii) If k is not employed, his expected productivity must be below $p^M(G(h_t))$, while $\mu^*_{h_t}$ guarantees employer j at least $p^M(G(h_t))v$. For blocking to strictly benefit employer j, worker k would need to pay j, which is not possible due to workers' limited liability.

Why don't flexible wages fix spiraling under breakdown learning? One plausible conjecture is that with flexible wages, workers with similar expected productivities obtain similar earnings. This would indeed be the case in the one-shot version of the model, because a worker with expected productivity p obtains flow earnings $\max\{0, (p - p^M)v\}$, which is indeed a continuous function of p. In particular, there would be no discontinuity in flow earnings between an unemployed worker ($p < p^M$) and a worker who barely makes the cut ($p \approx p^M$).

However, in a dynamic setting, employed workers benefit from the information that they generate through employment: unlike unemployed workers, they have an opportunity to establish an increasingly higher reputation and thereby command an increasingly higher wage, which quickly sets them apart from unemployed workers. To be sure, employed workers also risk generating a breakdown, in which case they become unemployed forever. However, such an event can occur only if a worker has a low type and, even in this case, it takes time to arrive, during which the low-type worker enjoys the benefits from employment. The accumulated learning for group a thus translates into a substantial earnings advantage over group b.

We now expand on this intuition about spiraling in two steps. First, to show how learning through employment strictly benefits a worker, consider a discretized version of the model with only two periods and in which $\alpha = \beta = 1$, as depicted in Figure 6. In the first period, a-workers and b-workers, who have comparable expected productivity, have comparable flow earnings: while only a-workers are hired, their wage is equal to 0 since p^M is equal to p_a. The performance of an a-worker in the first period splits the prior belief p_a into posterior beliefs 0 and $\overline{p}_a$. Since earnings are convex in beliefs, this splitting strictly benefits a-workers, whose expected flow earnings in the second period now equal w_2. Hence, first-period learning causes the gap in flow earnings to widen in the second period.

Second, even though the benefit from learning over each short period (i.e., over $[t, t + \mathrm{d}t)$) is small, this benefit accumulates over time. Because the posterior $\overline{p}_a$ is significantly more likely to occur than the zero posterior, the delay in employment experienced by b-workers does not vanish even as p_b gets arbitrarily close to p_a. By the time employers start hiring b-workers, they have already gained substantial knowledge about a-workers' types. Hence, the average flow earnings of a-workers are significantly higher than those of b-workers.

For breakthrough learning, by contrast, the delay in employment experienced by b-workers vanishes as $p_b \uparrow p_a$. Hence, a-workers do not get a chance to accumulate the benefit from employer learning. The contrast between environments persists.

Proposition 3.5. *Given matching μ^*, as $p_b \uparrow p_a$ the average lifetime earnings of a-workers converge to those of b-workers under breakthrough learning but not under breakdown learning.*

Persistent gap in employment, wage, and flow earnings under breakdown learning. Besides establishing that spiraling continues to arise with flexible wages, we further quantify the magnitude of such spiraling. Appendix C.2.3 derives and analyzes closed-form expressions for the employment rate, the average wage, and the average flow earnings of each group.

We first show that if task scarcity is sufficiently severe—in the sense that there are more high-type workers than tasks—the employment gap persists throughout workers' careers, even though it decreases over time (Proposition C.4). Owing to this nonvanishing delay in employment faced by group b, the wage gap is strictly increasing for a substantial amount of time. The wage gap starts shrinking only after employed b-workers accumulate enough learning, and shrinks to zero only in the limit $t \to \infty$ (Proposition C.3). See Figure 7 below for an illustration.[18]

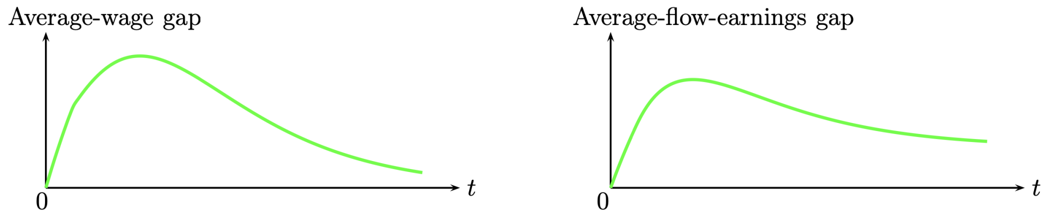


Figure 7: Average wage gap and average flow-earnings gap as $p_b \uparrow p_a$

The gap in flow earnings is due to the combination of the wage gap and the employment gap. Like the wage gap, the gap in flow earnings expands early in workers' careers and begins to gradually shrink only in the latter part of their careers (Proposition C.2). But unlike the wage gap, whether the gap in flow earnings shrinks to zero depends on how scarce the tasks are. If there are more high-type workers than tasks, this gap remains bounded away from zero even in the limit $t \to \infty$. This is due to the nonvanishing employment gap in the limit $t \to \infty$.

Spiraling persists even with negative wages. A natural question is whether spiraling would disappear if workers were able to accept negative wages in return for early employment opportunities. Appendix C.2.4 shows that the persistence of spiraling under flexible wages is not driven by the limited liability requirement, which we relax to a strictly negative lower bound on wages. Intuitively, since both a-workers and b-workers are willing to accept such negative wages, a-workers also lower their wages and outbid b-workers down to the lower bound. This intensifies competition among workers and benefits only the employers. The dynamic matching of workers and employers continues to be the same as that under limited liability, whereas all wages are now reduced by the same amount as the lower bound. In particular, the marginal-productivity worker at any instant now pays the employer the maximum amount that he can pay. As long as the lower bound is not too negative—which corresponds to b-workers having a strictly positive continuation payoff at $t = 0$—such a matching continues to be dynamically stable (Proposition C.5). Workers are willing to incur negative flow earnings so as to generate signals about their productivity. Therefore, b-workers continue to face the exact same delay as under limited liability.

4 Discussion and extensions

4.1 Empirical implications

Our analysis yields several testable implications about how inequality evolves in labor markets. Here, we begin by presenting four central predictions regarding how hiring, unemployment, and

[18]Figure 1 and Figure 7 assume the same parameter values: $\alpha = 5/4$, $\beta = 1$, $p_a = 1/2$, $\lambda_\ell = 1$, and $r = 1$.

wage dynamics differ across environments that emphasize breakdowns (visible failures) versus breakthroughs (visible successes). We then discuss how these predictions can be brought to the data via an occupation-level salience index that captures the relative visibility of failures and successes, connect our findings to the *Matthew effect*—the cumulative amplification of early advantage—and conclude with implications for affirmative action.

Prediction 1 (Early-career assignment and wage gaps). Among groups with comparable ex-ante productivity, breakdown-salient occupations display *larger early-career assignment gaps, longer unemployment spells, and wider wage gaps* than breakthrough-salient ones.

Our findings are consistent with documented gender disparities among physicians, aggregating across specialties: Lo Sasso et al. (2011) document a substantial unexplained pay gap for newly trained physicians, while Sarsons (2022) documents a referral gap disadvantaging female surgeons. Also, Lang and Lehmann (2012) observe that it is challenging to explain the simultaneous presence of large racial wage and employment gaps in low-skill occupations and the absence of such gaps in high-skill occupations. We provide a mechanism that can explain such discrepancies across occupations: to the extent that low-skill occupations are driven by breakdown learning and high-skill occupations by breakthrough learning, discriminated groups are likely to face more persistent wage gaps and longer unemployment spells in low-skill occupations.

Prediction 2 (Task scarcity and inequality). When job opportunities become *scarcer*—such as during recessions or hiring freezes—*inequality widens in breakdown-salient occupations.*

As the task-to-worker ratio falls, both groups lose, but the disadvantaged group loses more. At the aggregate level, therefore, greater task scarcity leads to more inequality. This helps explain why economic downturns—when labor demand falls—tend to deepen inequality. Estimates from the Pew Research Center show that the white-to-black and white-to-Hispanic wealth gaps peaked during the 2009 recession, the widest since data began in 1984.[19] Dahl and Knepper (2023) also document that age discrimination increases during times of higher unemployment.

Prediction 3 (Life-cycle dynamics). In breakdown-salient occupations, *wage gaps widen from early to mid-career and close only slowly thereafter; unemployment gaps persist; earnings gaps may never fully close.*

In breakdown-salient occupations, our analysis yields three related predictions. First, wage gaps widen from early to mid-career due to delayed opportunities for the disadvantaged group and only slowly narrow later in the career. Second, unemployment gaps can persist throughout the life cycle. Third, as a consequence, earnings gaps also widen from early to mid-career and may never fully close later in life because of the persistent unemployment disadvantage. These patterns are consistent with empirical evidence: Arcidiacono (2003) documents that the black–white earnings gap increases with experience and persists throughout the career, while Arcidiacono, Bayer and

[19] https://www.pewsocialtrends.org/2011/07/26/wealth-gaps-rise-to-record-highs-between-whites-blacks-hispanics/

Hizmo (2010) show that for high school graduates, ability is revealed only gradually and racial wage gaps that are small at entry widen with experience.

Prediction 4 (Speed and direction of employer learning). Faster employer learning *reduces but does not eliminate* inequality in breakdown-salient occupations.

In breakdown-salient environments, the dependence of the payoff gap on the arrival rate of breakdowns (in Proposition 2.2) implies that faster learning narrows—but never closes—the gap between comparable groups, even when low-productivity workers are identified almost instantaneously. Hence, at the aggregate level across occupations, accelerating employer learning alone (i.e., through policies that intensify monitoring or reporting without changing the direction of learning) may not resolve disparities. This underscores the joint importance of the direction and the speed of learning in shaping long-run inequality.

Figure 3 in Section 2.4 further illustrates this interaction, as shown by the blue solid curve. Holding the total arrival rate of signals fixed, payoff gaps tend to shrink with proximity to the corresponding pure environment—i.e., the salience of the environment. However, gaps are markedly more persistent on the breakdown side. Figure 3 also shows that, as groups' expected productivities diverge, breakthrough-salient and breakdown-salient environments start displaying similar payoff gaps that need not shrink with faster learning—so the direction and speed of learning become weaker predictors of inequality across groups.

Suggested operationalization of our framework

The taxonomy of occupations discussed in Section 2.1 offers one way to classify learning environments based on the skewness of empirical performance distributions. A complementary approach is to construct an occupation-level salience index that directly measures the relative visibility of failures and successes:

$$\text{Salience}_j = \frac{S_j^-}{S_j^- + S_j^+},$$

where S_j^- is the rate of major observable failures (malpractice claims, client losses, disciplinary actions) and S_j^+ is the rate of major observable successes (awards, renewals) in occupation j, with values near 1 indicating breakdown salience and near 0 indicating breakthrough salience.

Several complementary sources of variation can be used to test the model's predictions: *cross-occupation variation* (comparing industries differing in the observability of failures versus successes); *within-occupation variation* (exploiting institutional differences across firms or regions in how performance is recorded or disclosed); *temporal or cohort variation* (examining how shocks to the task-to-worker ratio change inequality across entry cohorts); and *variation in task scarcity* (using indicators such as procedures per physician or caseload per attorney to measure how task scarcity affects belief updating and wage inequality).

A cross-sectional implementation in medicine could leverage specialty-level data recording allocation of opportunities and resulting outcomes. Owda et al. (2025) report salary differences by gender, race, and ethnicity across 19 clinical specialties. Paired with specialty-specific salience measures, such data would allow a test of whether higher breakdown salience correlates with larger salary gaps. Sarsons (2022), while not focused on specialty-level disaggregation, provides detailed patient outcomes that proxy successes and failures as well as referral patterns that proxy allocation of opportunities. Together, these sources allow cross-specialty analyses of how success-vs.-failure visibility predicts inequality. Analogous constructions apply in law, finance, and academia, where wins/losses, sanctions/complaints, and recognition events are recorded and attributable.

Connection to the Matthew effect

The spiraling of small initial productivity differences into large and persistent payoff gaps is reminiscent of the *Matthew effect* (Merton, 1968), denoting cumulative advantage generated by early visibility and recognition. The term now encompasses a wide range of cumulative-advantage processes (Rigney, 2010). We contribute a learning-based mechanism through which the Matthew effect can arise and a classification of workplace environments by their susceptibility to it.

Merton (1968) emphasized reputation dynamics driven by visibility, skewed citation patterns, and institutional prestige. Subsequent research identified related channels: match-specific skill accumulation (Gibbons and Waldman, 1999), peer and network effects (Oyer, 2006), ability-sensitive technologies (Gabaix and Landier, 2008), confidence and self-perception (Murphy and Weinhardt, 2020), and workplace environments shaping fertility choices (Azmat, Cuñat and Henry, 2023). Our framework complements these by showing how the *direction of learning*—whether driven by failures or successes—can produce cumulative advantage even when underlying productivity differences are negligible.

A key insight in Merton (1968) is that cumulative advantage is amplified by the scarcity of recognition opportunities. In our framework, a similar scarcity—limited opportunities for employers to learn workers' productivity—drives spiraling in breakdown environments. The closest formal parallel is Bar-Isaac and Lévy (2022), who examine how task allocation affects signal generation; we instead emphasize how the *direction* of learning governs the magnitude of divergence.

Implications for affirmative action

A natural policy question is whether temporary affirmative-action initiatives can generate lasting effects ((Miller and Segal, 2012; Kurtulus, 2016; Miller, 2017)). The empirical evidence is mixed. In our framework, such interventions are especially effective in breakdown environments. A requirement that employers initially assign tasks to disadvantaged workers until expected productivities equalize can permanently narrow gaps: once beliefs converge, identical treatment follows even after the policy ends.[20] By contrast, in breakthrough environments—where success signals are self-correcting—the same intervention yields only transitory effects. This distinction clarifies when affirmative action acts as a catalyst for equality rather than a temporary correction.

[20] Such an intervention, though brief, is inherently fragile: if maintained beyond the period required to place the two groups on equal footing, it risks producing reverse discrimination.

4.2 Investment in productivity

If the workers had equal access to pre-market investment, would their lifetime employment prospects equalize? The answer depends on the investment equilibria that emerge. Access to investment could presumably level the playing field if incentives to invest were slightly stronger for group b, but it could alternatively amplify the expected productivity gap. We study this question in a variation of the small-market model with a three-step *pre-market investment stage*: (i) workers draw their pre-investment types independently of each other, according to the priors (p_a, p_b); (ii) a low-type worker of either group draws his cost of investment $c \in [0, 1]$ according to a distribution F and decides whether to invest; (iii) if he invests, he pays cost c and his type becomes high with probability $\pi \in (0, 1)$. A worker's investment cost, investment decision, and post-investment type remain private. Subsequently, workers enter the labor market at $t = 0$.

An equilibrium is characterized by a pair of cost thresholds (c_a, c_b) and a pair of post-investment beliefs about each worker's productivity (q_a, q_b). A low-type worker i invests if his realized cost is below c_i, and the employer's beliefs are consistent with this strategy. Denote by $B_i(q_a, q_b)$ worker i's expected benefit of investment given beliefs (q_a, q_b). Lemma 4.1 establishes that, in both the breakdown and the breakthrough learning environments, if the employer believes that worker i's expected productivity post-investment is higher than worker $-i$'s, then i is allocated the task first, so i's benefit from investment is strictly higher than $-i$'s.[21] However, a critical difference between the two environments is that this benefit function is continuous at $q_a = q_b$ in the breakthrough environment but not in the breakdown one.

Lemma 4.1. *In both pure 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 continuously differentiable in the pure breakthrough environment, but it is discontinuous at $q_a = q_b$ in the pure breakdown environment.*

The worker who is favored post-investment has a stronger incentive to invest, which in turn rationalizes the employer's decision to favor this worker in equilibrium. This self-fulfilling force—similar to that in Arrow (1973) and Coate and Loury (1993)—leads to multiple investment equilibria. We next explore the implications of Lemma 4.1 for inequality between groups across investment equilibria in either environment as $p_b \uparrow p_a$.

Investment as an equalizing force. The self-correcting property of breakthroughs remains undisturbed by the presence of investment. Proposition D.1 shows that as $p_b \uparrow p_a$, there exists an equilibrium in which the two workers' post-investment likelihoods of being a high type converge. The proof builds on two observations. First, when $p_a = p_b$, there always exists a symmetric equilibrium in which the workers use the same cost threshold and therefore $q_a = q_b$. Second, under breakthrough learning the benefit from investment is continuously differentiable in (q_a, q_b). By the implicit function theorem, when p_b is within a small neighborhood of p_a, there exists an

[21] In the breakthrough environment, if $q_a = q_b = q$, then $B_a(q, q) = B_b(q, q)$ because the employer optimally splits her task between the workers. The workers' benefits are also equal in the breakdown environment if the employer randomizes equally between workers with $q_a = q_b$ at $t = 0$.

equilibrium in which cost thresholds (c_a, c_b) and post-investment probabilities (q_a, q_b) are within a small neighborhood of those in the symmetric equilibrium. Although this equilibrium may preserve or reverse the workers' prior ranking, the resulting payoff gap remains small.

Additionally, the presence of investment can serve as an equalizing force in the breakdown environment. That is, there exists a mixed equilibrium in which $q_a = q_b$: the workers are post-investment identical due to worker b investing slightly more than worker a as $p_b \uparrow p_a$. This equilibrium relies on the employer randomizing *asymmetrically* at $t = 0$ between two post-investment equally productive workers, favoring worker b to provide him with slightly stronger investment incentives. Whether such an equilibrium is empirically plausible depends on the employer's ability to credibly and precisely randomize in this way.

Reversal of advantage and post-investment inequality. In either learning environment, investment can reverse the initial ranking of groups. If $(p_a - p_b)$ is sufficiently small, there exist equilibria in which worker b invests more than worker a and hence becomes favored post-investment. The presence of such reversal equilibria is certainly relevant for policy. Investment can end up not only equalizing the playing field, but overshooting and thus hurting the ex ante favored group.

To the extent that what matters is the payoff gap rather than the identity of the favored group per se, access to investment not only fails to tame the propensity of breakdown learning to magnify small prior differences, but can make it worse. Proposition D.2 shows that across all investment equilibria in which one worker is ahead post-investment, the expected payoffs of ex ante comparable workers are even further apart than in the no-investment benchmark of Section 2.1. This is due to the benefit from investment being discontinuous in (q_a, q_b). As $p_b \uparrow p_a$, there exist only two equilibria with a strict ex post ranking of workers and they are the same modulo the workers' identities. Inequality between workers increases due to investment. In the no-investment benchmark, the payoff ratio is pinned down by p_a. Here, because the worker who is favored post-investment has a strong enough incentive to invest, his post-investment probability is strictly higher than p_a. Therefore, for any realized investment cost, the ratio between the payoff of the worker who is discriminated against post-investment and that of the worker who is favored—after factoring in the investment cost—is lower than the ratio in the no-investment benchmark.

Breakdown environment preferred by the employer. Our characterization of equilibria allows us to compare learning environments not only in terms of the workers' payoffs, but also in terms of their investment behaviors. Proposition D.3 shows that with sufficiently fast learning, the worker favored (discriminated against) post-investment invests strictly more (less) under breakdowns than under breakthroughs. This ranking is robust across all investment equilibria in which one worker is ahead post-investment. Therefore, the breakdown environment is marked by greater polarization in workers' investment behavior. One key implication is that for sufficiently fast learning and effective investment ($\pi \approx 1$), the employer strictly prefers the breakdown environment to the breakthrough one, because the strong investment incentives provided by the breakdown environment guarantee that the post-investment favored worker is almost surely (in the limit as λ_ℓ, λ_h become arbitrarily large) a high type.

4.3 Inconclusive learning environments

Our baseline model assumed that signals were type-specific: only the high type can generate a breakthrough and only the low type can generate a breakdown. We now consider environments in which both types generate the same signals, albeit at different rates—hence, the arrival of a signal is inconclusive of the worker's type. The environment is characterized by a pair of arrival rates $(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$ such that $\lambda_\theta > 0$ is the arrival rate of the signal if the worker's type is $\theta \in \{h, \ell\}$. We define the environment to be an *inconclusive breakthrough environment* if the signal suggests a high type (i.e., $\lambda_h > \lambda_\ell > 0$) and an *inconclusive breakdown environment* otherwise (i.e., $0 < \lambda_h < \lambda_\ell$). If $\lambda_h = \lambda_\ell$, signals are uninformative.

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

Proposition 4.1. *For any $\lambda_h > \lambda_\ell$, the two workers' payoffs converge as $p_b \uparrow p_a$.*

The spiraling property generalizes to inconclusive breakdown environments as well. The departure from a conclusive breakdown environment brings the complication that the employer might hire workers who have generated signals 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 signal. The expected time until the first signal is significant. If players are sufficiently impatient or signals are sufficiently infrequent, this already creates a significant payoff advantage for worker a.

Proposition 4.2. *For any $\lambda_h < \lambda_\ell$, the two workers' payoffs do not converge as $p_b \uparrow p_a$ if*

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

The proof establishes a positive lower bound on the payoff gap by focusing on worker a's payoff until the arrival of the first signal and assuming continuation payoffs most favorable to worker b thereafter (i.e., continuation payoffs of 0 for a and 1 for b). This amounts to worker a's payoff before the first signal exceeding 0.5. To extend this argument, we use simulations to compute worker a's payoff before the first m signals: if it exceeds 0.5, spiraling necessarily arises. Figure 8 plots the region in the (λ_ℓ, r)-space where this holds for $m = 1$ (above the blue curve), $m = 2$ (above the red curve), and $m = 3$ (above the green curve). The region for $m = 1$ matches the condition in Proposition 4.2. As m increases, the spiraling region clearly expands. While simulations for larger m become computationally demanding, the outward shift suggests spiraling will hold across an even broader range of (λ_ℓ, r).

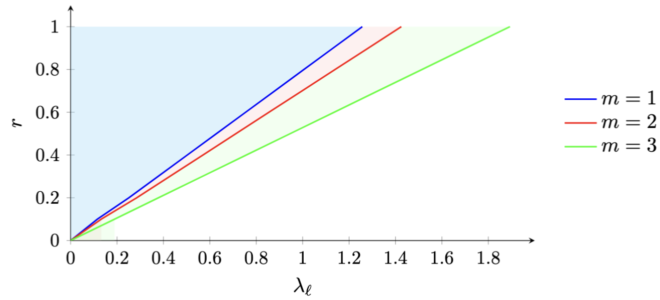


Figure 8: Simulated regions in the (λ_ℓ, r)-space for which spiraling is guaranteed to arise. Other parameters are $p_a = 1/3, p_b = 1/3 - 1/1000, \lambda_h = \lambda_\ell/2$, and $s = 0$.

4.4 Misspecified prior belief

Spiraling arises in the pure breakdown environment even if the groups are identical but the employer mistakenly misperceives them as different. Suppose the workers have the same probability p_{true} of having a high type, but the employer believes that worker b has a lower probability $p_{\text{mis}} < p_{\text{true}}$.[22] In this case, even a very slight misspecification grants a large payoff disadvantage to worker b. Worker a is still hired first based on the employer's misspecified belief and the workers' payoffs coincide with those in Proposition 2.2 (with p_a and p_b replaced by p_{true}).[23]

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 the workers. Duration t^*, which is analogous to the grace period in (4), corresponds to the time 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^*. At time t^*, the true probability that worker a has a high type is p_{mis}, whereas the true probability that worker b has a high type is p_{true}. However, the employer 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 worker b, respectively, at time t^*. Because each worker gets half a task but worker a has a lower true probability of having 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 2.1 to the misspecified-prior case, we let t^* be the time it takes for the belief to drop from p_{true} to p_{mis} and replace $U_i(p_b, p_b)$ with $\hat{U}_i(p_{\text{mis}}, p_{\text{true}})$ for the workers' payoffs.

Belief misspecification is highly relevant to 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.

[22] Bohren et al. (2023) refer to this as "inaccurate statistical discrimination." Bohren, Imas and Rosenberg (2019) identify discrimination driven by misspecified beliefs in an experimental setting.

[23] See payoff expressions (5) and (6) in Appendix B.

5 Concluding remarks

This paper studies the consequences of different employer learning environments for social groups of comparable expected productivity. 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 provides an additional perspective: what matters for statistical discrimination is not only the speed of employer learning, but also the direction of this learning.

Our framework may be used to address questions beyond the scope of the current paper. First, an employer may have to allocate multiple tasks that entail different employer learning dynamics. For instance, if an employer has both a breakthrough task and a breakdown task, how will she allocate the tasks among workers from comparable social groups? Second, in certain contexts the learning environment is an endogenous choice of the employer rather than being exogenously fixed. Our extension with productivity investment by workers identifies circumstances under which the employer prefers breakdown learning due to stronger investment incentives. More generally, is the endogenous choice of the learning environment more likely to lead to breakdown or breakthrough learning? If the employer can adjust her choice of the learning environment in response to the workers' expected productivities (as in Che and Mierendorff (2019)), how does this affect the lifetime payoffs of comparable groups? Third, our framework can prove useful for understanding incentives for occupational segregation: workers from discriminated groups have an incentive to sort into breakthrough-like occupations in order to avoid spiraling. We leave these questions to future research.

6 Data Availability Statement

The data and code underlying this research are available on Zenodo at https://doi.org/10.5281/zenodo.21072287.

References

Aguinis, Herman, and Kyle J. Bradley. 2015. "The Secret Sauce for Organizational Success." *Organizational Dynamics*, 44: 161–168.

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. 2026. "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.

Arcidiacono, Peter. 2003. "The Dynamic Implications of Search Discrimination." *Journal of Public Economics*, 87: 1681–1706.

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

Arrow, Kenneth. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets.* Vol. 3, 3–33. Princeton University Press.

Azmat, Ghazala, Vicente Cuñat, and Emeric Henry. 2023. "Gender Promotion Gaps: Career Aspirations and Workplace Discrimination." *Management Science*, Forthcoming.

Bar-Isaac, Heski, and Raphaël Lévy. 2022. "Motivating Employees Through Career Paths." *Journal of Labor Economics*, 40(1): 95–131.

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

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review*, 106(6): 1437–75.

Becker, Gary. 1957. *The Economics of Discrimination.* University of Chicago Press.

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. "Field Experiments on Discrimination." In *Handbook of Field Experiments.* Vol. 1, 309–393. 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(515): F480–F498.

Board, Simon, and Moritz Meyer-ter-Vehn. 2013. "Reputation for Quality." *Econometrica*, 81(6): 2381–2462.

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. 2023. "Inaccurate Statistical Discrimination: An Identification Problem." *Review of Economics and Statistics*, Forthcoming.

Bohren, J Aislinn, Peter Hull, and Alex Imas. 2025. "Systemic discrimination: Theory and measurement." *The Quarterly Journal of Economics*, 140(3): 1743–1799.

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

Bornmann, Lutz. 2024. "Skewed distributions of scientists' productivity: a research program for the empirical analysis." *Scientometrics*, 129(4): 2455–2468.

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

Che, Yeon-Koo, and Johannes Hörner. 2018. "Recommender Systems as Mechanisms for Social Learning." *The Quarterly Journal of Economics*, 133(2): 871–925.

Che, Yeon-Koo, and Konrad Mierendorff. 2019. "Optimal Dynamic Allocation of Attention." *American Economic Review*, 109(8): 2993–3029.

Che, Yeon-Koo, Kyungmin Kim, and Weijie Zhong. 2019. "Statistical Discrimination in Ratings-Guided Markets."

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.

Dahl, Gordon B, and Matthew Knepper. 2023. "Age Discrimination Across the Business Cycle." *American Economic Journal: Economic Policy*, 15(4): 75–112.

Deb, Rahul, Matthew Mitchell, and Mallesh M Pai. 2022. "(Bad) reputation in relational contracting." *Theoretical Economics*, 17(2): 763–800.

Drugov, Mikhail, Margaret Meyer, and Marc Möller. 2024. "Selecting the Best: The Persistent Effects of Luck."

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

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. 2021. "'Soft' Affirmative Action and Minority Recruitment." *American Economic Review: Insights*, 3(1): 1–18.

Flanagan, Robert J. 1978. "Discrimination Theory, Labor Turnover, and Racial Unemployment Differentials." *The Journal of Human Resources*, 13(2): 187–207.

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

Gabaix, Xavier, and Augustin Landier. 2008. "Why Has CEO Pay Increased so Much?" *The Quarterly Journal of Economics*, 123(1): 49–100.

Gibbons, Robert, and Michael Waldman. 1999. "A Theory of Wage and Promotion Dynamics Inside Firms." *The Quarterly Journal of Economics*, 114(4): 1321–1358.

Gittins, John, Kevin Glazebrook, and Richard Weber. 2011. *Multi-armed bandit allocation indices.* John Wiley & Sons.

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."

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

Halac, Marina, and Andrea Prat. 2016. "Managerial Attention and Worker Performance." *American Economic Review*, 106(10): 3104–32.

Halac, Marina, and Ilan Kremer. 2020. "Experimenting with Career Concerns." *American Economic Journal: Microeconomics*, 12(1): 260–88.

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.

Hurst, Erik, Yona Rubinstein, and Kazuatsu Shimizu. 2024. "Task-Based Discrimination." *American Economic Review*, 114(6): 1723–1768.

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.

Kelz, Rachel R., Morgan M. Sellers, Bijan A. Niknam, James E. Sharpe, Paul R. Rosenbaum, Alexander S. Hill, Hong Zhou, Lauren L. Hochman, Karl Y. Bilimoria, Kamal Itani, Patrick S. Romano, and Jeffrey H. Silber. 2021. "A National Comparison of Operative Outcomes of New and Experienced Surgeons." *Annals of Surgery*, 273(2): 280–288.

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

Komiyama, Junpei, and Shunya Noda. 2024. "On Statistical Discrimination as a Failure of Social Learning: A Multi-Armed Bandit Approach." *Management Science*, Forthcoming.

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.

Lepage, Louis-Pierre. 2023. "Experience-based Discrimination." *American Economic Journal: Applied Economics*, Forthcoming.

Li, Danielle, Lindsey R Raymond, and Peter Bergman. 2024. "Hiring as Exploration."

Lizzeri, Alessandro, Eran Shmaya, and Leeat Yariv. 2024. "Disentangling Exploration from Exploitation."

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.

Merton, Robert K. 1968. "The Matthew Effect in Science." *Science*, 159(3810): 56–63.

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.

Murphy, Richard, and Felix Weinhardt. 2020. "Top of the Class: The Importance of Ordinal Rank." *The Review of Economic Studies*, 87(6): 2777–2826.

Norman, Peter. 2003. "Statistical Discrimination and Efficiency." *The Review of Economic Studies*, 70(3): 615–627.

O'Boyle, Ernest, and Herman Aguinis. 2012. "The Best and the Rest: Revisiting the Norm of Normality of Individual Performance." *Personnel Psychology*, 65: 79–119.

Onuchic, Paula. 2023. "Recent Contributions to Theories of Discrimination."

Owda, D., M. O. Mensah, D. Yang, M. E. Canavan, C. P. Gross, and S. I. Chaudhry. 2025. "Salary Differences by Gender, Race, and Ethnicity Among Assistant Professors at US Medical Schools." *JAMA Network Open*, 8(5).

Oyer, Paul. 2006. "Initial Labor Market Conditions and Long-Term Outcomes for Economists." *Journal of Economic Perspectives*, 20(3): 143–160.

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

Pallais, Amanda. 2014. "Inefficient hiring in entry-level labor markets." *American Economic Review*, 104(11): 3565–99.

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

Rigney, Daniel. 2010. *The Matthew Effect: How Advantage Begets Further Advantage.* Columbia University Press.

Sarsons, Heather. 2022. "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.

Wallis, Christopher J. D., Bheeshma Ravi, Fahima Dossa, Nancy N. Baxter, Allan S. Detsky, Raj Satkunasivam, Carl van Walraven, David A. Urbach, Steven J. Hechtman, Robert S. Kaderbhai, Michael S. Boehler, Michael A. Coburn, H. Gilbert Welch, Allan S. Detsky, and Raj Satkunasivam. 2017. "Comparison of Postoperative Outcomes Among Patients Treated by Male and Female Surgeons: A Population-Based Matched Cohort Study." *BMJ*, 359: j4366.

A Derivation of the safe-arm threshold $\underline{p}$ for Section 2.1

Lemma A.1. *For any environment described by the pair of arrival rates $(\lambda_h, \lambda_\ell)$, the belief threshold at which the employer switches to the safe arm is given by*

$$\underline{p} := \frac{rs}{rv + \max\{\lambda_h, \lambda_\ell\}(v - s)}.$$

Proof of Lemma A.1. Consider first $\lambda_h > \lambda_\ell$. Fixing an arbitrary prior belief p and threshold belief $\underline{p} < p$, this corresponds to duration

$$t^*\left(p, \underline{p}\right) := \frac{1}{\lambda_h - \lambda_\ell} \log\left(\frac{p/(1-p)}{\underline{p}/(1-\underline{p})}\right).$$

Conditional on the worker having a high type, the payoff of the employer is

$$v\left(1 - e^{-rt^*(p,\underline{p})}\right) + \left(1 - e^{-\lambda_h t^*(p,\underline{p})}\right) e^{-rt^*(p,\underline{p})} v + e^{-\lambda_h t^*(p,\underline{p})} e^{-rt^*(p,\underline{p})} s,$$

whereas conditional on the worker having a low type, the employer's payoff is

$$\frac{\lambda_\ell + re^{-(\lambda_\ell + r)t^*(p,\underline{p})}}{\lambda_\ell + r} s$$

because the arrival probability of a breakdown at $t \leq t^*(p, \underline{p})$ is $\lambda_\ell e^{-\lambda_\ell t}$ and the employer's payoff is $e^{-rt}s$. Hence, the expected payoff of the employer simplifies to

$$V_{BT}(p) := pv - pe^{-(\lambda_h + r)t^*(p,\underline{p})}(v - s) + (1 - p)\frac{\lambda_\ell + re^{-(\lambda_\ell + r)t^*(p,\underline{p})}}{\lambda_\ell + r} s.$$

The smooth pasting condition yields

$$\lim_{p \downarrow \underline{p}} V'_{BT}(p) = 0 \quad \Rightarrow \quad \underline{p} = \frac{rs}{rv + \lambda_h(v - s)}.$$

Next, consider $\lambda_\ell > \lambda_h$. If the worker has a high type, the payoff of the employer is v: the worker is never fired, whether a breakthrough arrives or not. If the worker has a low type, the payoff of the employer equals the continuation payoff from the safe arm once a breakdown is realized, which is $\lambda_\ell s/(\lambda_\ell + r)$. Hence, the employer's expected payoff if she experiments with a worker given prior belief p is

$$V_{BD}(p) := pv + (1 - p)\frac{\lambda_\ell s}{\lambda_\ell + r}.$$

At the threshold $p = \underline{p}$, the employer is indifferent between the worker and the safe arm: the value matching condition is $V_{BD}(\underline{p}) = s$. This implies the threshold

$$\underline{p} = \frac{rs}{rv + \lambda_\ell(v - s)}.$$

■

B Proofs for Section 2

Proof of Proposition 2.1. The employer initially allocates the task exclusively to worker a. If worker a generates a breakdown, the employer switches to worker b. If worker a generates a breakthrough, worker a is hired forever. In the absence of either signal, this initial allocation lasts until the employer's belief that a has a high type decreases to p_b, which happens at time t^*, where t^* is defined by

$$\frac{p_a e^{-\lambda_h t^*}}{p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*}} = p_b, \quad \text{i.e.,} \quad t^* = \frac{1}{\lambda_h - \lambda_\ell} \log \frac{p_a(1-p_b)}{(1-p_a)p_b}. \tag{4}$$

If t^* is reached without a signal, the task is split equally between the two workers until either one of them generates a signal or the belief about the workers' types reaches $\underline{p}$. If worker i generates a breakthrough, the task is allocated only to him forever after. If worker i generates a breakdown, the task is allocated only to worker $-i$ until $-i$ generates a breakdown or $\underline{p}$ is reached.

We let $U_i(p_a, p_b)$ denote worker i's payoff given belief pair (p_a, p_b). Note that $U_a(p,p) = U_b(p,p) =: U(p)$ for any $p \in (\underline{p}, 1)$. Over $[0, t^*)$, worker a generates a breakthrough with probability $p_a\left(1 - e^{-\lambda_h t^*}\right)$ and a breakdown with probability $(1-p_a)\left(1 - e^{-\lambda_\ell t^*}\right)$. If a breakthrough arrives, worker a's payoff is 1. If a breakdown arrives at time t, worker a gets $(1 - e^{-rt})$. If no signal arrives, worker a's payoff consists of the flow payoff from $[0, t^*)$, which is $1 - e^{-rt^*}$, and the continuation payoff from time t^* onward, which is $U(p_b)$. Therefore, worker a's total expected payoff is

$$\begin{aligned} U_a(p_a, p_b) = p_a \left(1 - e^{-\lambda_h t^*}\right) + (1-p_a)\left(1 - e^{-\lambda_\ell t^*} - (1 - e^{-(\lambda_\ell + r)t^*})\frac{\lambda_\ell}{\lambda_\ell + r}\right) + \\ \left(p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*}\right)\left(1 - e^{-rt^*} + e^{-rt^*}U(p_b)\right). \end{aligned}$$

If worker a generates a breakdown before t^*, which occurs with instantaneous probability $(1-p_a)e^{-\lambda_\ell t}\lambda_\ell$, worker b gets discounted payoff $e^{-rt}K$, where K is the continuation payoff

$$\begin{aligned} K := p_b(1 - e^{-\lambda_h \underline{t}}) + (1-p_b)\left(1 - e^{-\lambda_\ell \underline{t}} - (1 - e^{-(\lambda_\ell + r)\underline{t}})\frac{\lambda_\ell}{\lambda_\ell + r}\right) \\ + (p_b e^{-\lambda_h \underline{t}} + (1-p_b)e^{-\lambda_\ell \underline{t}})(1 - e^{-r\underline{t}}) \end{aligned}$$

and $\underline{t} := 1/(\lambda_h - \lambda_\ell)\log\left((p_b(1-\underline{p}))/(\underline{p}(1-p_b))\right)$ is the time it takes for the belief to drop from p_b to the safe-arm threshold $\underline{p}$. On the other hand, if a generates a breakthrough over $[0, t^*)$, worker b gets zero payoff. Integrating over all $t \in [0, t^*)$, this payoff expression becomes

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(1 - e^{-(\lambda_\ell + r)t^*}\right)K.$$

If a does not generate any signals, which happens with probability $p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*}$, worker b gets $e^{-rt^*}U(p_b, p_b)$. Therefore, worker b's total expected payoff is

$$U_b(p_a, p_b) = (1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(1 - e^{-(\lambda_\ell + r)t^*}\right)K + (p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*})e^{-rt^*}U(p_b).$$

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

Proof of Proposition 2.2. We first observe that in any breakdown-salient environment, the expected payoff of each player depends on λ_ℓ but not λ_h. If worker a has a high type, his payoff is 1: because the belief about θ_a drifts upwards in the absence of a signal, worker a continues to be hired whether he generates a breakthrough or not. If he has a low type, his payoff is $(1 - e^{-rt})$ if the breakdown arrives at t, and this arrival time t follows density $\lambda_\ell e^{-\lambda_\ell t}$. Hence, worker a's expected payoff is

$$p_a + (1-p_a)\frac{r}{\lambda_\ell + r}, \tag{5}$$

which is independent of p_b. Similarly, if the employer starts hiring worker b at time t, then worker b's payoff is

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

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

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(p_b + (1-p_b)\frac{r}{\lambda_\ell + r}\right). \tag{6}$$

Evaluating the limit as $p_b \uparrow p_a$, worker b's expected payoff converges to

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(p_a + (1-p_a)\frac{r}{\lambda_\ell + r}\right),$$

which is equal to the fraction $(1-p_a)\lambda_\ell/(\lambda_\ell + r)$ of worker a's payoff. ■

Proof of Lemma 2.1. For any breakthrough-salient environment $\lambda_h > \lambda_\ell \geqslant 0$, the proof of Proposition 2.1 derives the payoff gap $U_a(p_a, p_b) - U_b(p_a, p_b)$, which does not depend on $U(p_b)$, the continuation payoff defined as in the proof of Proposition 2.1, and it simplifies to

$$\begin{aligned}
U_a(p_a, p_b) - U_b(p_a, p_b) &= p_a\left(1 - e^{-\lambda_h t^*}\right) + (1-p_a)\left(1 - e^{-\lambda_\ell t^*} - (1 - e^{-(\lambda_\ell + r)t^*})\frac{\lambda_\ell}{\lambda_\ell + r}\right) + \\
&\left(p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*}\right)\left(1 - e^{-rt^*}\right) - (1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(1 - e^{-(\lambda_\ell + r)t^*}\right)K \\
&= 1 - \left(p_a e^{-\lambda_h t^*} + (1-p_a)e^{-\lambda_\ell t^*}\right)e^{-rt^*} - (1-p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(1 - e^{-(\lambda_\ell + r)t^*}\right)(1+K),
\end{aligned}$$

where K, $\underline{t}$, and t^* are defined in the proof of Proposition 2.1. Because $\underline{p}, t^*$ and K all are continuously differentiable in $(\lambda_h, \lambda_\ell)$, this difference is continuously differentiable in $(\lambda_h, \lambda_\ell)$ as well. By a similar argument, the proof of Proposition 2.2 gives the payoff gap in any breakdown-salient environment, which after substituting in $\lambda_\ell = \lambda/2 - \delta$ simplifies to:

$$U_a(p_a, p_b) - U_b(p_a, p_b) = \frac{(-2\delta + \lambda)^2(p_a - p_b(1-p_a)) + 4(\lambda - 2\delta)p_a r + 4r^2}{(\lambda - 2\delta + 2r)^2}.$$

It is immediate that this is continuously differentiable in $\delta \in [-\lambda/2, 0)$. So it only remains to check that the limits of $U_a(p_a, p_b) - U_b(p_a, p_b)$ and of its derivative as $\delta \uparrow 0$ and $\delta \downarrow 0$ coincide.

Substituting $(\lambda_h, \lambda_\ell) = (\lambda/2 + \delta, \lambda/2 - \delta)$ into the payoff gap for a breakthrough-salient environment above and taking the limit $\delta \downarrow 0$, we obtain

$$U_a(p_a, p_b) - U_b(p_a, p_b) \to 1 - \frac{\lambda(1-p_a)(\lambda(1+p_b) + 4r)}{(\lambda + 2r)^2} = \frac{\lambda^2(p_a - p_b(1-p_a)) + 4\lambda p_a r + 4r^2}{(\lambda + 2r)^2}.$$

Differentiating this payoff gap with respect to δ and taking its limit as $\delta \downarrow 0$ gives us

$$\lim_{\delta \to 0} \quad \frac{\partial(U_a(p_a, p_b) - U_b(p_a, p_b))}{\partial \delta} = \frac{8(1-p_a)r(\lambda p_b + 2r)}{(\lambda + 2r)^2} > 0.$$

On the other hand, taking the limit of the payoff gap for the breakdown-salient environment above as $\delta \uparrow 0$, we obtain

$$U_a(p_a, p_b) - U_b(p_a, p_b) \to \frac{\lambda^2(p_a - p_b(1-p_a)) + 4\lambda p_a r + 4r^2}{(\lambda + 2r)^2}$$

which is exactly the limit obtained from the limit of a breakthrough-salient environment. Moreover, differentiating the breakdown-salient payoff gap with respect to δ and taking its limit as $\delta \uparrow 0$ gives $\frac{8(1-p_a)r(\lambda p_b + 2r)}{(\lambda + 2r)^2}$. Hence, the payoff gap is continuously differentiable in a neighborhood of $\delta = 0$. ■

Proof of Proposition 2.3. Consider first the class of breakdown-salient environments parametrized by λ. From the proof of Lemma 2.1, the payoff gap is

$$U_a(p_a, p_b) - U_b(p_a, p_b) = \frac{(-2\delta + \lambda)^2(p_a - p_b(1 - p_a)) + 4(\lambda - 2\delta)p_a r + 4r^2}{(\lambda - 2\delta + 2r)^2}$$

the derivative of which with respect to δ is

$$\frac{\partial(U_a(p_a, p_b) - U_b(p_a, p_b))}{\partial\delta} = \frac{8(1 - p_a)r(2r + \lambda p_b - 2\delta p_b)}{(2r + \lambda - 2\delta)^3} > 0$$

for $\delta \in [-\lambda/2, 0]$. By the continuity of the first derivative of the payoff gap established in Lemma 2.1, there exists a maximal $\tilde{\delta} \in (0, \lambda/2]$ such that the gap strictly increases in δ for $\delta \in (0, \tilde{\delta})$ as well.

■

C Proofs for Section 3

Consider a stage game. Let G denote the CDF of the distribution of the expected productivity p_k of worker $k \in [0, \alpha + \beta]$. Hence, $(\alpha + \beta)G(p)$ is the mass of workers with $p_k \leqslant p$. At time 0, p_k 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, more is learned about their types, causing G to evolve over time. The evolution of G depends, of course, on the learning environment.

We define $p^*(G)$ as the least productive worker within the most productive unit mass of workers. Throughout, we let $G(p^-) := \lim_{x \uparrow p} G(x)$ denote the left-hand limit of G at p.

Definition 3. Fix G. Let $p^*(G)$ be the expected productivity such that:

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

C.1 Proofs for Section 3.2 (Large market with fixed wages)

C.1.1 Stable stage-game matchings

We first establish that employers are matched to the most productive workers in any stable stage-game matching, provided that these workers are better than the safe arm. Let p_s be the belief at which a worker generates the same flow payoff to an employer as a safe arm does, that is, $p_s v - w = s$ where $w = 1$ is the fixed wage. Hence, $p_s = (s + w)/v$. We assume that $p_i > p_s$ for $i \in \{a, b\}$. Lemma C.1 shows that worker k is matched if $p_k > \max\{p^*(G), p_s\}$ and unmatched if $p_k < \max\{p^*(G), p_s\}$.

Lemma C.1 (Most productive workers are matched)**.** *Fix G and a stable stage-game matching (D, W). Let $d(p)$ denote the fraction of workers with an expected productivity of p who are matched to an employer.*

1. *Suppose that $p^*(G) > p_s$. Then $d(p)$ equals 1 if $p > p^*(G)$, and 0 if $p < p^*(G)$. If G is continuous at $p^*(G)$, then $d(p^*(G)) = 0$. If G is discontinuous at $p^*(G)$, then $d(p^*(G))$ is given by:*

$$(1 - G(p^*(G)))(\alpha + \beta) + d(p^*(G))\Big(G(p^*(G)) - G(p^*(G)^-)\Big)(\alpha + \beta) = 1.$$

2. *Suppose that $p^*(G) \leqslant p_s$. 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:*

$$(1 - G(p_s))(\alpha + \beta) + d(p_s)\big(G(p_s) - G(p_s^-)\big)(\alpha + \beta) \leqslant 1.$$

Proof. First, an employer never matches with worker k if $p_k < p_s$. Second, if a less productive worker is matched, then a more productive worker must be matched as well. By way of contradiction, suppose that for some $p_1 < p_2$, a p_1 worker is matched but a p_2 worker is not. The employer who is matched to the p_1 worker can form a blocking pair with the p_2 worker since $p_2 v - w > p_1 v - w$ and $w > 0$.

We now argue that, if $p^*(G) > p_s$, no employer takes the safe arm. Suppose otherwise. Then there exists an unmatched worker k with $p_k \geqslant p^*(G) > p_s$. Then an employer who is taking the safe arm can form a blocking pair with this worker k.

We next argue that, if $p^*(G) \leqslant p_s$, then all workers whose p is strictly above p_s must be matched. Suppose instead that some worker k is unmatched and $p_k > p_s$. The mass of workers whose p is weakly above p_k is strictly less than 1. Hence, there exists an employer who is either matched to a worker k' with $p_{k'} < p_k$ or taking the safe arm. This employer can form a blocking pair with the unmatched worker k. ■

Lemma C.1 characterizes the set of stable stage-game matchings for a given G. This set need not be a singleton. However, multiplicity arises only when $(1 - G(p_s))(\alpha + \beta) < 1$ and $G(p_s) - G(p_s^-) > 0$, as covered by part 2 of Lemma C.1. In this case, employers are indifferent between a positive mass of workers whose productivity is p_s and the safe arms. Because stable stage-game matchings are uniquely specified except in this case, we say that the stable stage-game matching is essentially unique.

C.1.2 Dynamic stability

In the dynamic setting, G evolves endogenously over time due to learning about workers' types. To argue that prescribing the essentially unique stable stage-game matching characterized in Lemma C.1 is dynamically stable, we show that conditions (i)-(iii) in Definition 2 are satisfied. Recall that $G(h_t)$ denotes the CDF of p_k at history h_t and $\mu^*(h_t)$ the essentially unique stable stage-game matching given $G(h_t)$, as characterized in Lemma C.1.

Proposition C.1 (Fixed wages)**.** *In both the pure breakthrough environment and the pure breakdown environment, μ^* is dynamically stable.*

Proof. (i) At h_t, a matched employer j's flow payoff is at least s, because $\max\{p^*(G(h_t)), p_s\} \geqslant p_s$. The distribution $G(h_{t+\mathrm{d}t})$, and hence j's continuation payoff from $t + \mathrm{d}t$ onwards, does not depend on j's deviation. Therefore, she does not strictly prefer to take a safe arm over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$.

(ii) Suppose that worker k is matched at history h_t according to $\mu^*(h_t)$. Let $p(t)$ denote this worker's expected productivity at history h_t. We focus on the case where $p(t) \in (0, 1)$, since if $p(t) = 1$, the worker will be matched forever, and if $p(t) = 0$, the worker will be unmatched forever. We next show that the worker 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 $\tau \geqslant t$ and suppose the worker is employed over $[t, \tau)$ as long as no breakdown arrives. Let $p(\tau)$ denote the worker's expected productivity at time τ conditional on no breakdown in $[t, \tau)$ and $OM(\tau)$ the expected amount of time that the worker is employed in $[t, \tau)$. Then,

$$OM(\tau) = \int_t^\tau \left(p(t) + (1 - p(t))e^{-\lambda_\ell(x-t)}\right) \mathrm{d}x = p(t)(\tau - t) + (1 - p(t))\frac{1 - e^{-\lambda_\ell(\tau-t)}}{\lambda_\ell}.$$

Expressing $p(\tau)$ in terms of $p(t)$ from

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

we obtain that $p(\tau)$ and $OM(\tau)$ satisfy the following condition:

$$OM(\tau) = \frac{p(\tau) - p(t)}{\lambda_\ell p(\tau)} + \frac{p(t)}{\lambda_\ell} \log\left(\frac{(1 - p(t))p(\tau)}{p(t)(1 - p(\tau))}\right).$$

It is readily verified that $OM(\tau)$ increases in $p(\tau)$. Staying unmatched over $[t, t + \mathrm{d}t)$ and then reverting to $\mu^*_{h_{t+\mathrm{d}t}}$ only makes $p(\tau)$, and thus $OM(\tau)$, lower than their respective values on path for any $\tau \geqslant t$. Therefore, the worker does not strictly prefer to stay unmatched for $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$.

(b) We next consider breakthrough learning. Pick any $\tau \geqslant t$. Let $\tilde{Q}(\tau)$ denote the probability that this worker generates a breakthrough in $[t, \tau)$, $p(\tau)$ the worker's expected productivity at time τ, conditional on no breakthrough over $[t, \tau)$, and $\tilde{OM}(\tau)$ the expected amount of time that the worker is employed over $[t, \tau)$, conditional on no breakthrough. Then, by similar calculations to those in part (a), we obtain that $p(\tau)$ and $\tilde{OM}(\tau)$ satisfy the following condition:

$$\tilde{OM}(\tau) = \frac{p(t) - p(\tau)}{\lambda_h(1 - p(\tau))} + \frac{p(t) - 1}{\lambda_h} \log\left(\frac{(1 - p(t))p(\tau)}{p(t)(1 - p(\tau))}\right).$$

By Bayes' rule,

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

It is readily verified that both $\tilde{OM}(\tau)$ and $\tilde{Q}(\tau)$ decrease in $p(\tau)$. The lower $p(\tau)$ is, the longer the worker has been employed over $[t, \tau)$ conditional on no breakthrough, and the higher the probability that the worker has generated a breakthrough over $[t, \tau)$. Staying unmatched over $[t, t + \mathrm{d}t)$ and then reverting to $\mu^*_{h_{t+\mathrm{d}t}}$ only increases $p(\tau)$ relative to its value on path, thus lowering both $\tilde{OM}(\tau)$ and $\tilde{Q}(\tau)$. This holds for every $\tau \geqslant t$, so the worker does not strictly prefer to stay unmatched for $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$.

(iii) Suppose otherwise that worker k and employer j are not matched to each other under $\mu^*(h_t)$ but both strictly prefer to be matched over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$. Under $\mu^*(h_t)$, any employer's flow payoff is at least $\max\{p^*(G(h_t)), p_s\}v - w$. If employer j finds it strictly preferable to match with k and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$, it must be that $p_k > \max\{p^*(G(h_t)), p_s\}$. This means that worker k was already matched under $\mu^*(h_t)$, so he must not strictly prefer to be matched with j since the wage is fixed.

■

Next, we use Lemma C.1 to fully characterize the dynamics of task allocation under each pure learning environment, given the evolution of the expected productivity distribution.

C.1.3 Breakthrough learning

Once a worker generates a breakthrough, he is employed for the rest of time. To track how many workers have "secured their jobs", we let $m(t) \in [0, 1]$ denote the mass of workers who have generated a breakthrough by t, so $(1 - m(t))$ 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 $\alpha > 1$ and $p_a > p_b$. 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 matched 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 and are allocated one full task thereafter. 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 $(\alpha - m(t))$ of a-workers have not generated a breakthrough. Each has a high type with probability $q(t)$, and is allocated $\frac{1-m(t)}{\alpha-m(t)} \in (0, 1)$ of a task. Therefore, the evolution of $m(t)$ follows:

$$\mathrm{d}m(t) = (\alpha - m(t))q(t)\lambda_h \frac{1 - m(t)}{\alpha - m(t)}\mathrm{d}t = q(t)\lambda_h(1 - m(t))\mathrm{d}t \quad \text{and} \quad m(0) = 0. \tag{7}$$

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

$$q(t)(\alpha - m(t)) + m(t) = p_a\alpha \quad \Longrightarrow \quad q(t) = \frac{\alpha p_a - m(t)}{\alpha - m(t)}. \tag{8}$$

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 $m(t)$ and $q(t)$ is similar. The evolution of $m(t)$ is the same as (7). By the law of large numbers, for any $t \geqslant T_b$, $q(t)$ satisfies:

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

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

Proof of Proposition 3.1. 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 (8), we have that

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

Therefore, as $p_b \uparrow p_a$, $m(T_b) \to 0$. Using the fact that (i) $m(0) = 0$, (ii) $m(t)$ is independent of p_b for $t < T_b$, and (iii) $m(t)$ is strictly increasing in t, we conclude that $T_b \to 0$.

Conditional on reaching T_b without a breakthrough, an a-worker has the same continuation payoff as a b-worker 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 $T_b \to 0$. ■

C.1.4 Breakdown learning

Under breakdown learning, a matched worker stays matched as long as no breakdown occurs. At $t = 0$, 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 matched before. This process goes on until all the a-workers are tried. From that instant onward, 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. Like in the case of breakthrough learning, this T_b is again the delay that is experienced by group b uniformly.

We let $\underline{m}(t) \geqslant 1$ be the mass of workers who have been tried before t. Among these workers, one unit are currently employed, and a mass $(\underline{m}(t) - 1)$ 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 $p_a\underline{m}(t)$, so $1 - p_a\underline{m}(t)$ are matched to low-type workers. Hence, the evolution of $\underline{m}(t)$ follows:

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

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

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

If $p_a\alpha < 1$, then T_b is finite and solves $\underline{m}(T_b) = \alpha$. 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 $p_a\alpha + p_b(\underline{m}(t) - \alpha)$. Hence, the evolution of $\underline{m}(t)$ follows:

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

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

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

We let T_s denote the time at which this process of hiring untried b-workers ends. If $p_a\alpha + p_b\beta < 1$, there are fewer high-type workers than employers. Therefore, the process of hiring untried b-workers ends when $\underline{m}(t)$ reaches $\alpha + \beta$. If $p_a\alpha + p_b\beta \geqslant 1$, there are weakly more high-type workers than employers, in which case the process of hiring untried b-workers never ends (so $T_s = \infty$). This is because learning becomes extremely slow when the mass of employers matched with low-type workers approaches zero.

Proof of Proposition 3.2. 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 an *i-worker's continuation payoff from the time he is first allocated the task.* From the proof of Proposition 2.2, we know that $V(p_i) = p_i + (1 - p_i)r/(\lambda_\ell + r)$. An a-worker's expected payoff is

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

A b-worker's expected payoff is

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

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 3.3 shows that the larger the labor force, i.e., the larger the mass of workers relative to the fixed unit mass of tasks, the greater the inequality across groups.

Proof of Proposition 3.3. 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 $\underline{m}(t)$ 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. ■