Early-Career Discrimination: Spiraling or Self-Correcting?

Arjada Bardhi Yingni Guo Bruno Strulovici*

April 25, 2023

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 track primarily 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 discrimination self-corrects so as to guarantee comparable lifetime earnings. This contrast is robust to large labor markets, flexible wages, inconclusive learning, investment in productivity, and misspecified employers' 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 the worker's gender or race—as statistical proxies for the worker's productivity.[1] Such early-career statistical discrimination determines who makes the first cut when job opportunities are scarce. Workers from social groups that are expected to be more productive may be systematically prioritized even if differences in the groups' productivity distributions are infinitesimally small.

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

*Bardhi: Department of Economics, Duke University; email: arjada.bardhi@duke.edu. Guo: Department of Economics, Northwestern University; email: yingni.guo@northwestern.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, Eric 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).

on-the-job performance and reallocate opportunities accordingly. However, an opposite conjecture suggests that comparable groups might 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 how employers learn about workers' productivity makes a critical difference for which conjecture is correct. In environments that track primarily on-the-job successes, early-career discrimination has only minor consequences for workers' later employment opportunities and lifetime earnings. In environments that track on-the-job failures, in contrast, early-career discrimination significantly affects workers' lifetime prospects. Moreover, its 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 when the impact of early-career discrimination vanishes or gets amplified over time.

We show that this contrast between learning environments persists even in large markets with flexible wages and negative minimum wages that allow workers to pay employers for early opportunities. Under flexible wage determination—a possibility that we formalize in a dynamic two-sided matching model with incomplete information—comparable groups face very different employment and wage paths in environments that track failures. Such environments introduce a substantial delay in employment for social groups that are slightly disfavored at the outset.

Model. Our analysis focuses on labor markets in which (i) workers from distinct groups compete for scarce tasks, (ii) employers learn about a worker's productivity only if the worker performs a task, and (iii) groups have comparable productivity distributions.[2] Sarsons (2019) studies one such market in which male and female surgeons compete for referrals from physicians. Physicians learn about surgeons' abilities from surgeries they have performed in the past. Sarsons (2019) 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 initial difference on workers' subsequent employment opportunities and earnings.

Scarcity of tasks relative to workers, which our model takes as exogenous, can arise from various factors. It can stem from increasing automation in the workplace, which reduces the demand for human labor. It can also arise from the pyramid structure of most organizations, as positions become scarcer at higher ranks. Other factors could be the high cost of setting up and maintaining certain job positions, as well as increasing specialization, which leads to only a few workers being needed in highly specialized positions.

Our analysis begins with a stylized *small-market model* that features one employer and two workers identified by their respective social groups, a and b. We then generalize the analysis to a

[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 provides a particularly stark illustration of the role played by the learning environment in the dynamics of statistical discrimination.

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

dynamic two-sided large market with a continuum of workers from each group and a continuum of employers. In the baseline model, a worker's productivity is either high or low, and worker a is ex ante more likely than worker b to have high productivity. At each instant, the employer allocates the task to one of the two workers—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 is increasing in the productivity of the employed worker, whereas workers benefit from being allocated the task regardless of their productivity.

The employer learns about a worker's productivity from observed performance. We contrast two learning environments: *breakthrough* learning and *breakdown* learning. In the breakthrough environment, a high-productivity worker generates successes, or "breakthroughs," at randomly distributed times, whereas a low-productivity worker generates no successes. In the breakdown environment, a low-productivity worker generates failures, or "breakdowns," at randomly distributed times, whereas a high-productivity worker generates no failures. In the context of the small market, we analyze mixed learning environments that combine both breakthroughs from high-productivity workers and breakdowns from low-productivity workers; the relative frequency of the signals determines whether the environment is *breakthrough-salient* or *breakdown-salient.*[4]

The employer's 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). In terms of performance, star jobs have a high upside and a low downside, while guardian jobs have a low upside and a high downside. High-stakes salespeople, entertainers, and athletes are examples of star jobs, while routine surgeons, airline pilots, and prison guards are examples of guardian jobs.[5] (Figure 5 in appendix A.1 compares the performance distributions of star jobs and guardian jobs.[6])

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's. 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 in the two workers' expected productivity 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

[4]In such mixed 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-productivity (resp., high-productivity) workers also generate breakthroughs (resp., breakdowns) but at a lower frequency than high-productivity (resp., low-productivity) workers.

[5]Bose and Lang (2017) argue that most nonmanagerial jobs are guardian jobs and derive the optimal monitoring policy for such jobs. We instead compare the lifetime impact of early-career discrimination in star jobs to that in guardian jobs.

[6]Similarly to the performance distribution of star jobs in Figure 5, O'Boyle and Aguinis (2012) and Aguinis and Bradley (2015) show that in occupations centered around star performance, such as entertainers and athletes, 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).

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 employers becomes arbitrarily fast*. It can explain why societies struggle to eliminate inequality in labor markets.

This contrast between these two pure learning environments extends to mixed environments with both conclusive breakthroughs and conclusive breakdowns, even though the dynamic allocation of tasks is now more involved. Any breakthrough-salient environment—in which successes are revealed more often than failures—continues to be self-correcting, whereas any breakdown-salient environment gives rise to spiraling. However, the contrast is also more nuanced: fixing the total rate at which a worker's type is revealed but varying the arrival rate of breakthroughs relative to breakdowns, the payoff gap changes smoothly as the environment shifts continuously from a pure breakthrough environment to a pure breakdown one. In fact, the payoff gap is single-peaked over such a parametrized family of environments, with a peak that is arbitrarily close to the symmetric environment in which both types are revealed equally fast. The more comparable groups are, the more pronounced is the contrast between the two classes of environments: the payoff gap is close to zero for any breakthrough-salient environment and close to the strictly positive gap of the pure breakthrough environment in any breakdown-salient environment.

We further explore this contrast in a large market with many workers from each group and many employers. The key determinant of the spiraling effect in the breakdown environment is the scarcity of tasks relative to the size of each group. As tasks become scarcer, the inequality between groups increases. This implies that, while all groups suffer from a decrease in labor demand during economic downturns, *groups that are discriminated against will suffer disproportionately more*.

One might a priori conjecture that wage flexibility eliminates inequality between groups with comparable productivity. For instance, Becker (1957) and later Flanagan (1978) have argued that if employers can flexibly engage in wage discrimination, the wage differential should equalize the employment rates across groups. To evaluate this conjecture, we introduce flexible wages into the large market described in the previous paragraph. 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 (2020)).

We find that wage flexibility does not resolve the severe differential treatment of comparable groups in the breakdown environment. Intuitively, flexible wages do not overcome the tension caused by relative task scarcity: when only a subset of workers can be hired, those who generate higher surplus in expectation get hired first. Hence, as in the case of fixed wages, workers from group b experience a delay in employment relative to workers from group a: b-workers are not given a chance to perform unless and until sufficiently many a-workers have experienced breakdowns. This 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 wage and earnings gaps between groups of almost identical expected productivity.

Figure 1 illustrates the predicted paths of average wages and those of average earnings for two such groups.[7] Both the gap in average wages and that in average earnings expand in the early part of workers' careers, and persist for a substantial amount of time. If tasks are sufficiently scarce relative to workers, *the earnings gap persists throughout workers' careers.* While our main characterization assumes a zero minimum wage, we show that the dynamics are similar *with a strictly negative minimum wage* that affords workers to compensate employers for early opportunities to prove their productivity to the market. Workers outbid each other to the point that the marginal-productivity worker hired at each instant is paid the negative minimum wage. However, the employment paths are identical to those under a zero minimum wage, so b-workers face the same delay.

The spiraling of a negligible productivity difference into a substantial payoff gap is reminiscent of the cumulative advantage known as the Matthew effect (Merton, 1968). Scarcity of opportunities for acknowledgement is at the core of both our argument and that in Merton (1968). While the Matthew effect has become an umbrella term for cumulative advantage resulting from various mechanisms, our findings uncover a novel learning-based mechanism for this effect and identify workplace environments that are more prone to it.

We further demonstrate the robustness of the contrast between learning environments to workers strategically investing in their productivity, to inconclusive performance signals, and to prior differences being due to employers' incorrect beliefs. In particular, when workers can invest in their productivity before entering the labor market, the contrast becomes even sharper. Across all investment equilibria of the breakdown environment, slightly different groups invest in significantly different amounts. Inequality across groups is even greater (i.e., spiraling is even worse) than in the baseline model, since access to investment disproportionately benefits the group favored post-investment. In the breakthrough environment, in contrast, there generically exists an equilibrium in which comparable groups invest in comparable amounts and obtain comparable lifetime earnings. Hence, the self-correcting property of the breakthrough environment persists with investment opportunities.

Empirical implications and evidence. Our findings are consistent with the persisting gender pay gap among surgeons documented by Lo Sasso et al. (2011) and Sarsons (2019). In line with our emphasis on early-career discrimination, a recent statement by the Association of Women Surgeons finds that "[T]he disparities women face in compensation at entry level positions lead to a persistent trend of unequal pay for equal work throughout the course of their careers."[8] Our

[7] Average 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 earnings.

[8] For more, see the 2017 Association of Women Surgeons' Statement on Gender Salary Equity

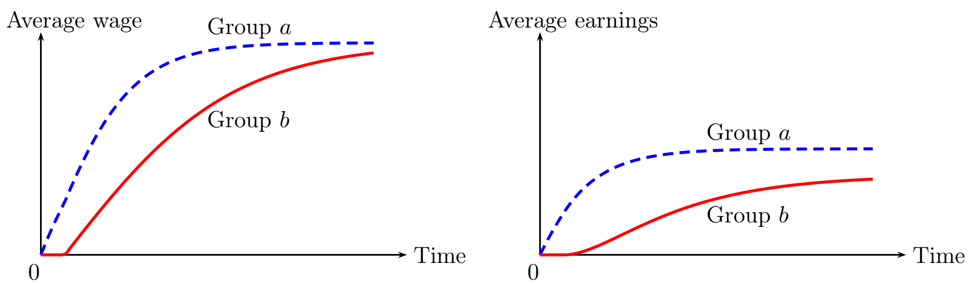


Figure 1: Average wage/earnings under breakdowns for groups of comparable productivity

results are also consistent with empirical evidence of racial wage gaps that are small at early career stages but widen with labor market experience, as documented by Arcidiacono (2003) and Arcidiacono, Bayer and Hizmo (2010). We provide a learning-based mechanism that can explain this growing wage gap across groups.

In contrast to the breakdown environment, the paths of employment rates, average wages, and average earnings in the breakthrough environment are arbitrarily close across the two groups. 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. Our model provides a mechanism that can explain such discrepancies across occupations. We predict that, all else being equal, breakthrough-like occupations tend to exhibit smaller and more transient wage and employment gaps than breakdown-like occupations. To the extent that low-skill occupations tend to be breakdown environments and high-skill occupations tend to be breakthrough ones, we provide an explanation for the more persistent wage gaps and longer unemployment duration faced by groups discriminated against in low-skill occupations.

Lastly, prejudice can be another cause of early-career discrimination: even when different groups have the same productivity distribution, employers may mistakenly believe that one group's distribution is inferior to the other's. Such prejudice may be caused by inaccurate stereotypes or inaccurate information about the workforce that enters a particular occupation. The contrast between breakthrough and breakdown environments extends to this setting as well, as we show in section 4.4. In a breakdown environment, prejudice among employers, even if very mild, can have dire consequences for the group that is discriminated against.

Related literature

First and foremost, our paper contributes to the theoretical literature on statistical discrimination surveyed in Fang and Moro (2011). Phelps (1972) and the subsequent literature (e.g., Aigner and Cain (1977), Cornell and Welch (1996), and Fershtman and Pavan (2020)) 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

(https://www.womensurgeons.org/page/SalaryStatement).

difference between groups: inequality arises because groups coordinate on different equilibria or specialize in different roles within an equilibrium.[9]

Our approach differs from both of these strands of literature. First, we consider groups that share arbitrarily similar expected 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 repeated task allocation.

Our analysis also contributes two insights to the literature on *cumulative discrimination* (e.g., Blank, Dabady and Citro (2004), Blank (2005)). First, the nature of the employer learning environment has a critical impact on the magnitude of cumulative discrimination. Second, the prospect of future cumulative discrimination casts a long shadow on workers' investment in productivity. Similarly to Pallais (2014), our findings emphasize the informational value of entry-level jobs.

Our results also speak to the literature on employer learning (e.g., Farber and Gibbons (1996), Altonji and Pierret (2001)). The learning environment can be interpreted as an intrinsic feature of an occupation. In this respect, our work relates to Altonji (2005), Lange (2007), Antonovics and Golan (2012), and Mansour (2012). Whereas these models assume that occupations differ only in the frequency of signals, we allow the direction of these signals to differ across occupations and demonstrate the importance of such direction.

Our analysis 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)).[10] Both Felli and Harris (1996) and this paper use the framework of multi-armed bandits to model labor market learning.[11] In our setting, the employer is the bandit operator and the workers are the bandit arms. The contrasting of breakthrough and breakdown learning environments 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),

[9] 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.

[10] 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)).

[11] In an extension which explores the workers' incentives to invest in productivity, the quality of the bandit arms is endogenously determined. By modeling bandit arms as strategic players, this extension is related to Bergemann and Valimaki (1996), Felli and Harris (1996), and Deb, Mitchell and Pai (2020). However all these models assume that the quality of the bandit arms is exogenously given.

Che and Hörner (2018), and more recently Lizzeri, Shmaya and Yariv (2021).

There is a growing literature on bandit problems and statistical discrimination. Li, Raymond and Bergman (2020) designs a screening algorithm that values exploration and thus leads to higher quality and diversity of interviewed candidates. Lepage (2020) assumes that employers are uncertain about minority groups' productivity. He shows that early negative signals from minority workers deter future learning and lead to group differential in the long run. Komiyama and Noda (2020) examines social learning by short-run employers, showing that population imbalance can lead to under-exploration of minority groups. Che, Kim and Zhong (2019) also considers a social learning model: even with identical groups, there exist discriminatory equilibria in which one group remains under-explored. Fershtman and Pavan (2020) shows that policies that aim to recruit more minority candidates can backfire if evaluation of minority groups is noisier.

2 Small market

2.1 Framework

Players and types. We consider a dynamic task allocation game in a small labor market consisting of 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 $p_i \in (0, 1)$. The employer knows (p_a, p_b), but she does not observe the workers' types. We interchangeably refer to p_i as the *prior belief* for worker i and 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. The worker obtains a flow payoff $w > 0$ whenever he is assigned the task. Otherwise, his flow payoff is zero. The parameter w, which is normalized without loss to $w = 1$, can be interpreted as the fixed wage for a worker who performs the task. 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 is paid. If the employer allocates the task to the safe arm, she earns a flow payoff $s \in (0, v)$. The employer's payoffs are observed at the end of the horizon.[12]

[12] We scale players' lifetime payoffs by a factor of r, as in Keller, Rady, and Cripps (2005). This normalizes the employer's lifetime payoff from a high type to v and a worker's lifetime payoff from being allocated the task for the entire horizon $[0, \infty)$ to 1.

Learning by allocating tasks. Learning about a worker's type proceeds via conclusive Poisson signals.[13] 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 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. A breakthrough perfectly reveals a high type and a breakdown perfectly reveals a low type. Thus, a *learning environment* is characterized by a pair of type-dependent arrival rates $(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$.

Central to our analysis are the pure learning environments in which only one signal is possible. In a pure breakthrough environment, $\lambda_h > 0$ and $\lambda_\ell = 0$, whereas in a pure breakdown one, $\lambda_h = 0$ and $\lambda_\ell > 0$. A mixed learning environment combines the two pure environments by featuring both conclusive signals: breakthroughs from the high type and breakdowns from the low type.[14] We distinguish between two classes of learning environments based on which of the types is revealed more quickly:

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

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

The relative frequency of each signal determines the evolution of the employer's belief in the absence of a signal. In a breakthrough-salient (resp., breakdown-salient) environment, the employer becomes more pessimistic (resp., optimistic) about the worker's type. In the symmetric environment with $\lambda_h = \lambda_\ell$, such absence is entirely uninformative of the worker's type. We interpret the learning environment as an intrinsic feature of how performance is monitored and evaluated at a given job. Breakthroughs corresponds to over-performance by high-type workers, and breakdowns to under-performance by low-type workers. Therefore, breakthrough-salient environments aim at identifying star employees and breakdown-salient ones aim at identifying misfits.

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

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

The belief 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.

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

[14] Appendix D.2 extends our framework to allow for inconclusive signals in the pure environments. Despite some loss in tractability, we establish the self-correcting property in Proposition 4.3 and the spiraling property for sufficiently impatient players in Proposition 4.4, discussed in section 4.

2.2 Contrast between the pure learning environments

We first observe a stark contrast between the workers' lifetime payoffs in the two pure environments, and then formalize and generalize this contrast in section 2.3. We compare the two workers' payoffs when the expected productivity of one group is arbitrarily close to that of the other and analyze how this comparison depends on whether the signal is a breakthrough or a breakdown.

Pure breakthrough learning. At each instant, the employer allocates the task to the worker with the higher expected productivity.[15] In the absence of a breakthrough the employer becomes more pessimistic about the worker's type. 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 for as long as no breakthrough arrives, worker a is effectively given a grace period $[0, t^*)$ over which to perform, 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 the belief drops down to $\underline{p}$, so the workers obtain the same continuation payoff starting from t^*.

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

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 allocates the task to the worker with the higher expected productivity. As long as this worker generates no breakdown, the employer becomes more optimistic that his type is high. She first allocates the task to worker a and continues to hire him in the absence of a breakdown. 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.

Therefore, 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, he faces an average delay of $\lambda/(\lambda + r)$ until worker a's low type is revealed. Crucially, this delay is independent of how close p_b is to p_a: even if p_b is just slightly less than p_a, worker a

[15] This is true in any learning environment that we consider because (i) workers have binary types, and (ii) the arrival rates of signals and the employer's type-contingent flow payoffs are the same for both workers.

obtains a substantially higher payoff than worker b. 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.[16]

Remark 1 (Continuum of productivity types). The contrast across learning environments goes beyond binary productivity types. For example, suppose that worker i's productivity type is uniformly distributed $\theta_i \sim U[\underline{\theta}_i, \underline{\theta}_i + \Delta]$, where $\underline{\theta}_a > \underline{\theta}_b > 0$ and $\Delta > 0$ parametrizes the uncertainty about the workers' types. As $\underline{\theta}_b \uparrow \underline{\theta}_a$, the expected productivities of the two groups converge. A signal that perfectly reveals the worker's type arrives according to a Poisson process, and its arrival rate is type-dependent. If $\lambda(\theta)$ increases (or decreases) in θ—e.g., $\lambda(\theta) = \lambda\theta$ (or $\lambda(\theta) = \lambda/\theta$) for some $\lambda > 0$,—the learning environment is a breakthrough (or breakdown) one and the employer becomes more pessimistic (or optimistic) in the absence of a signal. The self-correcting property of breakthroughs and the spiraling property of breakdowns continue to hold. In particular, the upward belief drift in the absence of a breakdown gives the first worker a payoff advantage that does not depend on how far the prior distributions are.

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. An infinitesimal belief difference leads to an infinitesimal lifetime payoff difference across workers in breakthrough-salient environments but to a substantial payoff gap in breakdown-salient environments. While this limit observation suggests a discontinuity of the payoff gap across environments, the payoff gap is in fact continuous in the arrival rates $(\lambda_h, \lambda_\ell)$ for any small but positive prior belief difference. This allows us to analyze how the payoff gap changes as the the learning environment shifts from a breakdown-salient environment to a breakthrough-salient one for a small prior difference.

[16] 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.

Self-correcting property of breakthrough-salient learning. In such an environment, both types can be revealed conclusively but the high type is revealed more frequently than the low type. Therefore, 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 effectively given a grace period $[0, t^*)$ over which to perform as long as his type is not revealed to be low, where

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

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. Naturally, t^* is proportional to the rate of learning $\lambda_h - \lambda_\ell$. 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 immediately hires worker b instead. Otherwise, starting from t^*, the employer splits the task equally between the workers until the belief drops down to $\underline{p}$, so the workers obtain the same continuation payoff starting from t^*. As in the pure breakthrough case, as $p_b \uparrow p_a$, the grace period t^* shrinks to zero and so does worker a's advantage. Therefore, any breakthrough-salient environment gives rise to self-correction in the belief limit.

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. In a breakdown-salient environment, the low type is revealed more frequently than the high type, so 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. On the other hand, if a breakdown is realized, the employer turns to worker b immediately. In the absence of either signal, the employer continues to hire worker a. Similarly to the pure breakdown case, this guarantees that a high-type worker is never fired, whereas a low-type worker enjoys a substantial period of employment before being eventually fired. Because 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 players' payoffs and the resulting payoff gap in any breakdown-salient environment are the same as in the corresponding 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), which is the same as that observed in the pure breakdown environment, 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. Moreover, 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 two types are revealed with the same frequency, so the absence of a signal is entirely uninformative. Worker a is fired if and only if a breakdown arrives revealing his low type. On the other hand, worker b obtains an opportunity, with some delay, only if worker a has a low type. The task allocation dynamics and the workers' payoffs in this environment with symmetric environment are therefore the same as in any breakdown-salient environment with the same arrival rate of breakdowns λ_ℓ. The limit payoff ratio is that in (3) and spiraling arises.

2.4 Comparative statics with respect to the learning environment

Our analysis so far fixes arrival rates $(\lambda_h, \lambda_\ell)$—that is, it fixes the learning environment—and examines the limit payoff gap as p_b approaches p_a. This section performs the complementary exercise of fixing expected productivities $p_a > p_b$ and examines how the payoff gap 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 encodes the full spectrum of learning environments 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 $(\lambda_h, \lambda_\ell) = (\lambda, 0)$.

Taken together, Propositions 2.1 and 2.2 imply a sharp discontinuity in the neighborhood of the symmetric environment at the limit $p_b \uparrow p_a$: as $\lambda_h \uparrow \lambda/2$, the limit payoff ratio for $p_b \uparrow p_a$ is strictly less than 1, whereas as $\lambda_h \downarrow \lambda/2$ the limit payoff ratio is 1. That is, an environment that is slightly breakdown-salient generates much higher inequality between workers than one that is slightly breakthrough-salient. However, this discontinuity goes away if we repeat the exercise for any positive belief difference $p_a - p_b > 0$, which is tractable since the proofs of Propositions 2.1 and 2.2 derive a general expression for the gap in the workers' expected payoffs $U_a(p_a, p_b) - U_b(p_a, p_b)$ for any (p_a, p_b) and $(\lambda_h, \lambda_\ell)$. For any $p_a > p_b$, the payoff gap is in fact continuously differentiable in the arrival rates, including in a small neighborhood around the symmetric environment.

Lemma 2.1. *Fix $p_a > p_b > \underline{p}$ and $\lambda > 0$, and consider a family of environments $(\lambda_h, \lambda_\ell) = (\lambda/2 + \epsilon, \lambda/2 - \epsilon)$ where $\epsilon \in [-\lambda/2, \lambda/2]$ and $\lambda > 0$. The payoff gap is continuously differentiable in ϵ.*

Figure 2 illustrates the payoff gap for various differences in prior beliefs. Although for any given learning environment the payoff gap naturally decreases as the belief difference gets smaller, the contrast between breakthrough-salient environments and breakdown-salient ones becomes more pronounced as well. Strikingly, the plot illustrates that the payoff gap is single-peaked across the family of environments parametrized by λ, with a peak reached at a breakthrough-salient environment. As $p_a - p_b$ gets arbitrarily small, the environment that generates the largest payoff gap approaches the symmetric environment. The higher $|\epsilon|$ is, the more informative the absence of a signal is for the employer and the more quickly she updates her belief about the worker. Therefore, although faster learning within each class of environments tends to ameliorate the inequality between comparable workers, this is not the case for breakthrough-salient environments that are sufficiently close to the symmetric environment.

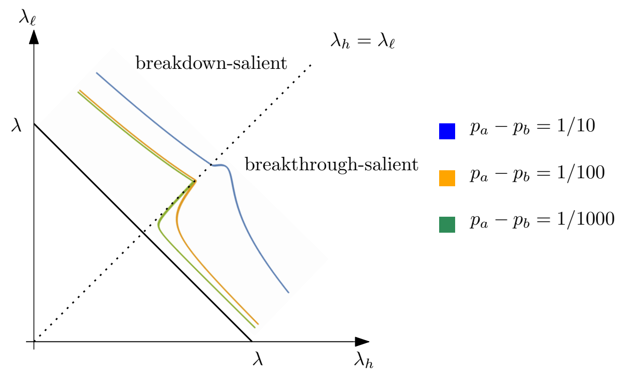


Figure 2: The payoff gap is illustrated for three different levels of the prior belief gap for environments parametrized by $\lambda_h + \lambda_\ell = \lambda$. The payoff gap is the height of the colored curve above the black curve. Parameter values: $\lambda = 5$, $s = 1/10$, $v = 1$, $r = 1$, $p_b = 1/3$, $p_a = p_b + \epsilon$ where $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}\}$.

Proposition 2.3 formalizes these observations about the monotonicity of the payoff gap. First, the gap unequivocally decreases the closer the environment is to a pure breakdown environment within the class of breakdown-salient environments. Faster revelation of low types decreases the delay faced by worker b, which always benefits worker b much more than it hurts a low-type worker a. Second, within the class of breakthrough-salient environments, the payoff gap increases as the environment becomes more breakthrough-salient for environments sufficiently close to the symmetric one. In such an environment, faster learning disproportionately benefits a high-type worker a, who faces a deadline to generate a breakthrough, thus contributing to a higher gap. In contrast, in more pronounced breakthrough-salient environments, the marginal benefit of faster learning—i.e., higher revelation of a high type and slower revelation of a low type—is higher for worker b, which shrinks the payoff gap.

Proposition 2.3. *Consider a family of environments* $(\lambda_h, \lambda_\ell) = (\lambda/2 + \epsilon, \lambda/2 - \epsilon)$ *where* $\epsilon \in [-\lambda/2, \lambda/2]$ *and* $\lambda > 0$.

(i) For any $p_a > p_b > \underline{p}$*, the payoff gap strictly increases in* ϵ *for* $\epsilon < \tilde{\epsilon}$*, where* $\tilde{\epsilon} > 0$.

(ii) Fixing $p_a > \underline{p}$*, there exist* $\eta \in (0, p_a - \underline{p})$ *and* $\bar{\epsilon} \in (\tilde{\epsilon}, \lambda/2)$ *such that for any* $p_b \in (p_a - \eta, p_a)$ *and any* $\epsilon \in (\bar{\epsilon}, \lambda/2]$*, the payoff gap* $U_a(p_a, p_b) - U_b(p_a, p_b)$ *strictly decreases in* ϵ.

Example 1. If the prior expected productivities are sufficiently far apart, the pure breakthrough environment can in fact lead to a larger payoff gap than the pure breakdown one. To see this, suppose $r = 1, s = 1/10, v = 1, p_b = 1/3, \lambda = 5$ as in the parametrization of Figure 2, and let $p_a = 2/3$. The payoff gap in the pure breakthrough environment is ≈ 0.621, whereas that in the pure breakdown environment is ≈ 0.598. Therefore, the contrast between the two pure environments that was observed in section 2.2 can be reversed if the groups are far from comparable.

3 Large market

The small-market model of section 2 featured a single employer choosing between only two workers. We now turn to a large two-sided market with a continuum of employers and a continuum of workers from each group. Employers and workers are matched dynamically, tasks are scarce, and all learning is public. We establish that the contrast between breakthrough environments and breakdown environments not only generalizes to this large market, but it does so irrespectively of whether wages are fixed or flexible. In particular, this means that flexible wages are insufficient to avoid spiraling in the breakdown environment. Strikingly, this is true even if the minimum wage is strictly negative, in which case the workers would be able to pay employers for early opportunities. Section 3.2 studies the dynamics of task allocation under fixed wages, whereas section 3.3 examines the case of flexible wages, both with a zero minimum wage and a strictly negative one.

3.1 Framework

There is a unit mass of employers, a mass of size $\alpha > 1$ of a-workers, and a mass of size β of b-workers.[17] Crucially, there are more workers than tasks in this large market.[18] All employers and workers are long-lived and share the same discount rate $r > 0$. Employers are ex ante homogeneous. At $t = 0$, each worker's type is drawn independently from other workers' types. The type of an i-worker is high with probability p_i, where $i = a, b$. There is also a unit mass of identical safe arms available. At each instant, each employer has one task to allocate and each worker can take up at most one task.

Stable stage-game matchings. In this dynamic market, a stage-game matching specifies (i) how workers are matched to employers, and (ii) a wage for each matched pair. Section 3.2 imposes that wages are fixed and constant across all matched pairs, as in the baseline model, whereas section 3.3 allows for flexible wages.

Specifically, we use $i \in [0, \alpha + \beta]$ to index a worker and $j \in [0, 1]$ to index an employer. Worker i is from group a if $i \in [0, \alpha]$ and from group b if $i \in (\alpha, \alpha + \beta]$. In the stage game, let $D_{ij} \in \{1, 0\}$ indicate whether worker i and employer j are matched to each other. If they are, let W_{ij} denote their wage, which is $W_{ij} = 1$ in the case of fixed wages and $W_{ij} \geqslant 0$ in the case of flexible wages. If $D_{ij} = 1$, worker i's payoff is W_{ij} and employer j's payoff is $p_i v - W_{ij}$, where p_i denotes the employer's belief that $\theta_i = h$. All signals are observed publicly, so all employers share the same belief about each worker. If $D_{ij} = 0$ for all j, worker i is unmatched and gets zero payoff. If $D_{ij} = 0$ for all i, employer j takes a safe arm and gets a payoff $s > 0$, which in the stage game corresponds to the belief threshold $p_s := s/v$. We assume that an employer's flow payoff from an a-worker or a b-worker given prior beliefs p_a and p_b, respectively, is higher than that from the safe

[17] We continue to model only two groups for ease of exposition. The results can be readily extended to more than two groups.

[18] The assumption that there are more a-workers than tasks simplifies exposition since only a-workers are matched at $t = 0$. However, our results hold qualitatively for $\alpha < 1$ as well, as long as tasks are relatively scarce, i.e., $\alpha + \beta > 1$.

arm: $p_s < p_b < p_a$. As before, we seek to understand the limit case $p_b \uparrow p_a$. Let $\mathcal{D}$ be the set of all such stage-game matchings.

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

Definition 1. A stage-game matching (D, W) is *stable* if (i) there exists no employer j matched to some i who strictly prefers to take a safe arm instead, and (ii) there exists no blocking pair.

Dynamic stability. In the dynamic game, let $\mathcal{H} := \cup_{t \geqslant 0} \mathcal{H}_t$ be the set of all histories and $\mathcal{H}_t$ the set of all time-t histories. A time-t history consists of all past matchings and realized signals until t. A *dynamic matching* $\mu = (\mu_t)_{t \geqslant 0}$ specifies a lottery over stage-game matchings for any history, i.e., $\mu_t : \mathcal{H}_t \to \Delta(\mathcal{D})$ for each t. We define dynamic stability of a dynamic matching based on the solution concept of a stable convention in Ali and Liu (2020).[19] For a given dynamic matching μ, let $\mu|_h$ denote the continuation matching after some history h.

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

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

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

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

In what follows, sections 3.2 and 3.3 show that, irrespective of whether wages are determined flexibly, prescribing the essentially unique stable stage-game matching after each history is dynamically stable: no worker-employer pair and no single player has a profitable one-shot deviation after any history.

3.2 Fixed wages

The intuition obtained from the small market of section 2 extends to this large market: comparable groups continue to have comparable payoffs in breakthrough environments but markedly different payoffs in breakdown ones. Moreover, the large market clarifies the role of task scarcity in the size of the payoff gap. The scarcer tasks are relative to the workforce, the greater the inequality between groups in the breakdown environment.

Diverse hiring under breakthrough learning. Mirroring the analysis in section 2.2, the dynamic allocation of tasks goes through two phases. In the first phase, all a-workers take turns to perform tasks. If an a-worker generates a breakthrough, he "secures his job" with his current

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

employer: the employer allocates future tasks only to this worker. For those a-workers without a breakthrough, the employers' belief drops gradually until it reaches p_b. At that point, a-workers without breakthroughs are believed to be as productive as b-workers. The allocation now enters a second phase in which the remaining employers let all remaining a-workers and all b-workers take turns to perform tasks. Again, those who generate breakthroughs secure their jobs with their current employers.

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

> For a star job, the costs of a hiring error are small relative to the upside potential from finding an exceptional individual. Therefore, the organization will wish to sample widely among many employees, looking for the one pearl among the pebbles. (Baron and Kreps (1999), p. 28-29)

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

Proposition 3.1 (Self-correction in a large market)**.** *Let $\lambda_h > \lambda_\ell = 0$. For $\alpha > 1$ and $\beta > 0$, the expected payoff of an a-worker converges to that of a b-worker as $p_b \uparrow p_a$.*

Narrow hiring under breakdown learning. Breakdown learning, in contrast, leads to sluggishness in experimenting with new workers: if a worker is hired, he remains employed until he generates a breakdown. This sluggishness hurts group b disproportionately no matter how close p_b is to p_a, thus generalizing the intuition behind Observation 2.2 to larger labor markets.

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

Proposition 3.2 (Spiraling in a large market)**.** *Let $\lambda_\ell > \lambda_h = 0$. 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 and 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. We measure group inequality by the ratio of a b-worker's expected payoff to that of an a-worker. Proposition 3.3 observes that in the breakdown environment, inequality between groups increases as the size of either group increases while the mass of tasks is kept fixed.

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

Proposition 3.3 (Inequality increases in task scarcity under breakdown learning)**.** *Let $\alpha > 1$ and $\beta > 0$. 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.*

This result predicts that when the scarcity of job opportunities intensifies, e.g., when labor demand falls during an economic downturn, inequality deepens. This is consistent with the observation that while all groups suffer during an economic downturn, some suffer disproportionately more.[20]

3.3 Flexible wages

This section incorporates flexible wages into the dynamic large market introduced above. In particular, it investigates whether flexible wages can restore earnings equality in the breakdown environment.[21] We show that, as long as workers' wages are bounded below by limited liability, both the self-correcting property of breakthroughs and the spiraling property of breakdowns continue to hold. Wage flexibility is insufficient to prevent spiraling.

Appendix C.2.1 characterizes the set of stable stage-game matchings. On the equilibrium path, there is a time-dependent marginal productivity $p^M(t)$ such that, at each time t, workers whose expected productivity (i.e., whose probability of having a high type) exceeds $p^M(t)$ are matched and workers whose expected productivity lies below $p^M(t)$ are not. Wages take a strikingly simple form: a matched worker with expected productivity p_t at time t is paid a flow wage of $(p_t - p^M(t))v$, which is the additional value that he creates relative to the marginal-productivity worker. An unmatched worker receives no wage and, hence, has zero earnings. All employers get the same flow profit of $p^M(t)v$.[22]

Proposition C.2 establishes that prescribing the stable stage-game matching after each history is dynamically stable. First, in a stable stage-game matching, an employer's flow profit from a match is at least as high as that from the safe arm, so no employer finds it profitable to reject a

[20] Estimates from the Pew Research Center (https://www.pewsocialtrends.org/2011/07/26/wealth-gaps-rise-to-record-highs-between-whites-blacks-hispanics/) show that the white-to-black and white-to-Hispanic wealth ratios were much higher at the peak of the recession in 2009 than they had been since 1984, the first year for which the U.S. Census Bureau published wealth estimates by race and ethnicity based on the Survey of Income and Program Participation.

[21] We assume that workers do not know their types at time 0: they share the same prior belief as the employers and all learning is symmetric. This assumption—which simplifies the stability approach taken in this section—is standard in models of learning in labor markets, such as Felli and Harris (1996) and Altonji and Pierret (2001).

[22] For notational ease, in this section $(v, 0)$ denotes the employer's *gross* flow payoffs from high and low types, respectively (i.e., prior to paying the flexible wage to the worker), whereas in the rest of the paper $(v, 0)$ denotes her *net* flow payoffs (i.e., after the fixed unit wage is paid).

match and take the safe arm. Second, no employer-worker pair has a profitable one-shot deviation, since all employers make the same flow profit. Lastly, one can show that no worker ever finds it profitable to reject a match in the hope of delaying the arrival of information about his type. This last point follows from the fact that a worker's flow earnings are convex in his expected productivity p_t at time t: flow earnings take the form of $\max\{0, (p_t - p^M(t))v\}$, as Figure 3 shows.[23] By Jensen's inequality, this implies that any signal about the worker's type at time t—which splits the current belief about the worker's type into a lottery over posterior beliefs—increases the worker's earnings, in expectation, at *all* future dates.

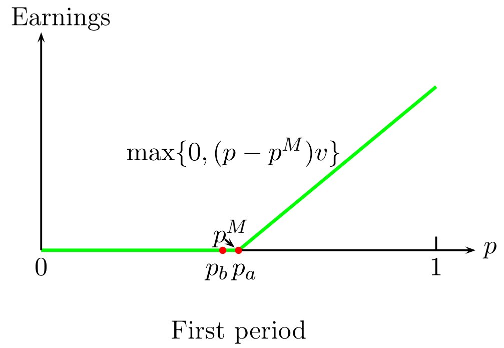


Figure 3: A two-period example with $\alpha = \beta = 1$

Flexible wages do not fix spiraling under breakdown learning. One plausible conjecture is that with flexible wages, workers with similar expected productivity 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 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. Naturally, they may also generate a breakdown and become unemployed. However, such an event occurs only if a worker has a low type and even then it takes time. The accumulated learning for the favored group thus translates into a substantial earnings advantage over the less favored group.

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 3. In the first period, a-workers

[23] The minimum wage for an employed worker is zero since we normalize workers' cost of effort on the task to zero. If this cost were strictly positive, the limited liability constraint would require that the wage be weakly greater than this cost. In the left panel of Figure 1 the average-wage paths for both groups would shift up by this cost, whereas in the right panel the average-earnings paths would remain intact once reinterpreted as average-*net*-earnings paths. Moreover, the green curve in Figure 3 would be reinterpreted as *net* flow earnings.

and b-workers, who have comparable expected productivity, have comparable 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 earnings in the second period now equal w_2. Hence, first-period learning causes the earnings gap 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, such benefit accumulates over time. Because $\overline{p}_a$ is significantly more frequent 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 that employers start hiring b-workers, they have already learned a lot about a-workers' types. Hence, the average earnings of a-workers are significantly higher than those of b-workers.

For breakthrough learning, in 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 average earnings of both groups thus converge.

Persistent employment, wage, and earnings gaps under breakdown learning. Besides establishing the fact that spiraling continues to arise with flexible wages, we are able to quantify the magnitude of such spiraling. Appendix C.2.3 computes and analyzes closed-form expressions for the employment rate, the average wage, and the average 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.6). 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 sufficiently many b-workers have been employed, and shrinks to zero only in the limit $t \to \infty$ (Proposition C.5). See figure 4 below for an illustration.[24]

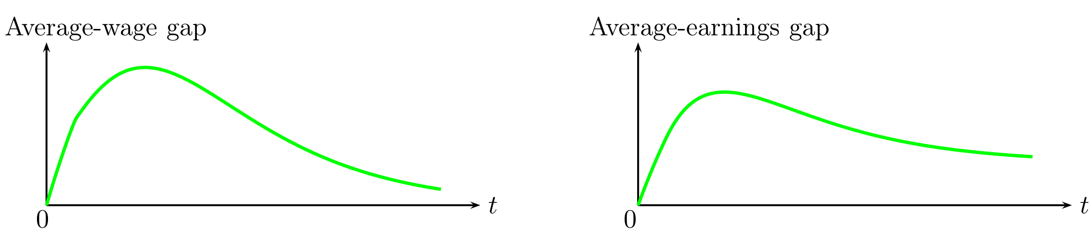


Figure 4: Average-wage gap and average-earnings gap as $p_b \uparrow p_a$

The earnings gap is due to the combination of the wage gap and the employment gap. Like the wage gap, the earnings gap expands early in workers' careers and begins to gradually shrink only in the latter part of their careers (Proposition C.4). But unlike the wage gap, whether the earnings gap also shrinks to zero depends on how scarce the tasks are. If there are more high-type

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

workers than tasks, the earnings 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 b-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 equally 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. We show that as long as the lower bound is not too negative—which corresponds to b-workers having a strictly positive continuation value at $t = 0$—such a matching continues to be dynamically stable (Proposition C.7). 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 robustness

4.1 Connection to the Matthew effect

The spiraling of a negligible productivity difference into a substantial payoff gap is reminiscent of the cumulative advantage known as the Matthew effect (Merton, 1968). In coining the term, Merton (1968) observed that scientists of established reputation tend to receive a disproportionately larger share of the credit for joint and simultaneous discoveries, which advances their reputation further. He observed that the effect is intrinsically linked to scarcity of opportunities for acknowledgment, which he referred to as the phenomenon of the "41st chair". Scarcity of opportunities for workers to prove their productivity is what drives spiraling in breakdown environments as well.

The Matthew effect has become an umbrella term for cumulative advantage that results from various mechanisms in science and beyond (Rigney (2010)). Merton (1968) observed that in science, reputation buildup could be due to greater visibility in the scientific community, skewed citation patterns, and institutional prestige and resources. A recent literature in economics explores various such mechanisms.[25] We contribute a novel learning-based mechanism through which the Matthew effect could arise, as well as a classification of workplace environments into more and less prone to this effect. Most closely related to us is Bar-Isaac and Lévy (2022), where task allocation also provides workers with opportunities to generate signals about productivity. However,

[25]For instance, the Matthew effect could arise due to the development of match-specific skills in labor markets in Gibbons and Waldman (1999), due to peer effects and differential institutional resources in Oyer (2006), due to the sensitivity of the production technology to the worker's ability in Gabaix and Landier (2008), due to heightened confidence from relative performance in Murphy and Weinhardt (2020), due to the friendliness of the workplace environments and fertility choices in Azmat, Cuñat and Henry (2020) etc.

whereas their focus is on the relationship between signal informativeness and worker's effort, ours is on the importance of the *direction* of employer learning for whether the Matthew effect arises.

4.2 Investment in productivity

If the workers had equal access to an opportunity to invest in productivity before entering the labor market, would their lifetime employment prospects equalize? The answer depends on the equilibrium implications of this investment opportunity. Access to investment could presumably level the playing field if incentives to invest are stronger for group b, or otherwise it could amplify the expected productivity gap. We study this question in a variation of the small-market model with a pre-market investment stage, which involves three steps: (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 cumulative distribution function 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 are observed only by this worker. 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). Worker i invests if the realized cost is below c_i and the employer's beliefs are consistent with this investment strategy. A key object in the analysis is worker i's expected benefit $B_i(q_a, q_b)$ from investment given the employer's post-investment belief pair (q_a, q_b). Lemma 4.1 establishes that in both pure learning environments, if the employer believes that worker i's expected productivity post-investment is higher than worker $-i$'s, then i's benefit from investment is strictly higher than $-i$'s.[26] Worker i would be the first to be allocated the task: investment is likely to avoid a breakdown in the near future or increase the chance of a breakthrough within the given grace period.

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—also noted in Coate and Loury (1993)—leads to multiple investment equilibria. In fact, 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 becomes favored post-investment. However, our focus is on the inequality across groups rather than the identity of the favored group per se. We characterize the lowest payoff inequality attained across all equilibria as $p_b \uparrow p_a$ in each learning environment.

[26] 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 equal also in the breakdown environment, assuming that the employer randomizes equally between workers at $t = 0$ if $q_a = q_b$.

Proposition 4.1 establishes that the lowest payoff inequality continues to be zero in the pure breakthrough environment. The self-correcting property of breakthroughs is not disturbed by the presence of investment. 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). We apply the implicit function theorem to establish that, when p_b is within a small neighborhood of p_a, there exists an equilibrium in which cost thresholds (c_a, c_b) and post-investment probabilities (q_a, q_b) are within a small neighborhood of those in the symmetric equilibrium. This equilibrium could either preserve or reverse the prior ranking of the workers.

Proposition 4.1. *Suppose that F is weakly convex. For a generic set of parameters,*[27] *as $p_b \uparrow p_a$, there exists an equilibrium in which the two workers' expected payoffs as well as their post-investment probabilities of having a high type converge.*

In contrast, access to investment not only fails to tame the propensity of pure breakdown learning to magnify small prior differences, but makes it worse. Across all investment equilibria, the expected payoffs of ex ante comparable workers are even further apart than in the no-investment benchmark of section 2.1.

Proposition 4.2 (Exacerbated spiraling in the presence of investment)**.** *As $p_b \uparrow p_a$, in any pure-strategy equilibrium (q_i, q_{-i}): (i) $q_i > q_{-i}$, and (ii) the ratio of the expected payoff of worker $-i$ to that of worker i is at most $(1 - q_i)\lambda_\ell/(\lambda_\ell + r) < 1$, which is strictly lower than the payoff ratio $(1 - p_a)\lambda_\ell/(\lambda_\ell + r)$ in the no-investment benchmark.*

Spiraling persists because the benefit from investment is discontinuous in (q_a, q_b). We show that 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—whoever that might be—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 to that of the worker who is favored—after factoring in the investment cost—is lower than the ratio in the no-investment benchmark. There exists also an equilibrium in which $q_a = q_b$—the workers are post-investment identical—due to worker b investing slightly more than worker a. This equilibrium relies on the employer randomizing between workers at $t = 0$ so as to provide slightly stronger investment incentives for worker b. Whether such an equilibrium is empirically plausible depends on the employer's ability to credibly and precisely randomize in this way.

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 behavior. Proposition D.1 in appendix

[27]The notion of genericity here is that fixing all parameters of the model except for (p_a, π), the set of values of $(p_a, \pi) \in (\underline{p}, 1) \times (0, 1)$ for which the proposition does not hold has measure zero. If F is not weakly convex, our preliminary analysis suggests that a version of this result continues to hold according to a different, more involved notion of genericity based on prevalent and shy sets.

D.1 shows that with sufficiently fast learning, the worker favored (discriminated) post-investment invests strictly more (less) under breakdowns than under breakthroughs. This ranking is robust across all investment equilibria. 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, since the strong investment incentives provided by the breakdown environment guarantee that the the post-investment favored worker is almost surely a high type.

4.3 Inconclusive learning environments

The baseline model assumed that signals are conclusive: in a mixed learning environment as introduced in section 2.1, only the high type can generate a breakthrough and only a low type can generate a breakdown. However, an even more general learning environment would allow for both types to generate the same signals, albeit at different rates—hence, the arrival of the signal would be informative but inconclusive of the worker's type. This extension analyzes such a learning environment with a single signal. 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 breakthrough, there is still a time window $[0, t^*)$ over which worker a should generate a first breakthrough in order to continue being allocated the task exclusively. If no breakthrough arrives during this time window, the belief about worker a's type drops to p_b, at which point both workers receive 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 breakthrough within the time window vanishes as well. The two workers' limit payoffs are therefore equal.

Proposition 4.3 (Self-correcting property of inconclusive breakthroughs)**.** *For any $\lambda_h > \lambda_\ell$, the two workers' payoffs converge as $p_a \downarrow p_b$.*

The spiraling property generalizes to inconclusive breakdown environments as well, provided that players are sufficiently impatient. The departure from a conclusive breakdown environment brings the complication that the employer might reconsider hiring workers who have generated breakdowns in the past. But as long as $p_a > p_b$, worker a is the first to be hired and stays employed in the absence of a breakdown. The expected time until the first breakdown is significant. If players are sufficiently impatient, this already leads to a significant payoff advantage for worker a.[28] Proposition 4.4 presents the details.

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

Proposition 4.4 (Spiraling property of inconclusive breakdowns)**.** *For any $\lambda_h < \lambda_\ell$, the two workers' payoffs do not converge if $r^2 - (1-2p_a)r(\lambda_\ell - \lambda_h) - \lambda_h\lambda_\ell > 0$ or equivalently:*

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

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 not identical. 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 $p_{\text{mis}} < p_{\text{true}}$.[29] 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}).[30]

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 very 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.

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

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

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

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 nature of that learning.

Our analysis has implications for how negative shocks to labor demand during economic downturns impact inequality across groups. We predict that breakdown-like occupations will be prone to significant increases in inequality as jobs become scarcer. To the extent that low-skill occupations tend to be predominantly breakdown environments and high-skill occupations tend to be breakthrough ones, our result is in line with substantial evidence that the groups who are hit the hardest in recessions are those who are already discriminated against and in low-skill occupations. Moreover, our results provide a learning-based explanation for the empirical observation that racial wage gaps are more present in low-skill occupations, which are typically breakdown-like, but are largely absent in high-skill ones (Lang and Lehmann (2012)). By this same reasoning, we explain why wage gaps might even widen with labor market experience in low-skill occupations, as documented by Arcidiacono, Bayer and Hizmo (2010). Our theoretical framework—and in particular, our predictions for the employment gap, the wage gap, and the earnings gap—can guide future empirical investigation of discrimination in breakthrough versus breakdown occupations.

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

Finally, our framework can be used to study questions that fall beyond the scope of the current paper. First, an employer may have to allocate multiple tasks which 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 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 to understanding incentives for occupational segregation: workers from groups that are discriminated against have an incentive to sort into breakthrough-like occupations in order to avoid spiraling. We leave these questions to future research.

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. 2020. "Conventions and Coalitions in Repeated Games."

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

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

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

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. 2020. "Gender Promotion Gaps: Career Aspirations and Workplace Discrimination." *Available at SSRN 3518420*.

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.

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. 2019. "Inaccurate Statistical Discrimination."

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

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

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." *Working paper.*

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

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

Deb, Rahul, Matthew Mitchell, and Mallesh M. Pai. 2020. "(Bad) Reputation in Relational Contracting." *Theoretical Economics*, Forthcoming.

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. 2020. "'Soft' Affirmative Action and Minority Recruitment."

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.

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.

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

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

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

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

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

Komiyama, Junpei, and Shunya Noda. 2020. "On Statistical Discrimination as a Failure of Social Learning: A Multi-Armed Bandit Approach." *Working paper.*

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. 2020. "Endogenous Learning, Persistent Employer Biases, and Discrimination." *Working paper.*

Li, Danielle, Lindsey R Raymond, and Peter Bergman. 2020. "Hiring as Exploration." National Bureau of Economic Research.

Lizzeri, Alessandro, Eran Shmaya, and Leeat Yariv. 2021. "Disentangling Exploration from Exploitation." 22nd ACM Conference on Economics and Computation. Keynote Lecture.

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.

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.

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

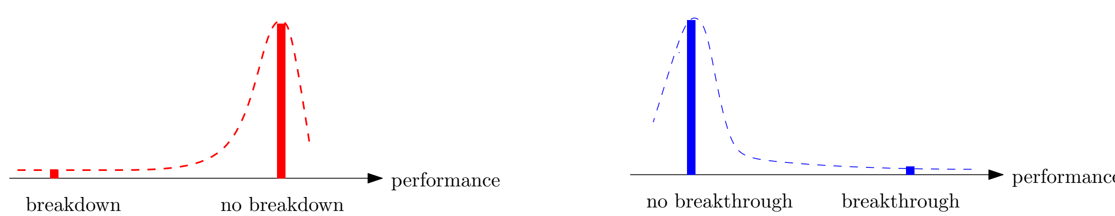
Shapley, Lloyd S, and Martin Shubik. 1971. "The Assignment Game I: The Core." *International Journal of Game Theory*, 1(1): 111–130.

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

A Preliminary results

A.1 Distribution of performance signals for star and guardian jobs

Replicating Figure 2-2 in Baron and Kreps (1999), the dashed curves in Figure 5 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. 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 5a) and conditional on a high type under breakthroughs (Figure 5b). The figure suggests how to empirically test whether a job is a star (breakthrough-like) job or a guardian (breakdown-like) one: a right-skewed density suggests a star job while a left-skewed density suggest a guardian job. See footnote 6 for examples of such empirical studies.



(a) Guardian jobs (b) Star jobs

Figure 5: Distribution of performance signals (adapted from Baron and Kreps (1999))

A.2 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

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

The smooth pasting condition yields

$$\frac{\partial V_{BT}(p, \underline{p})}{\partial \underline{p}} = 0 \quad \Rightarrow \quad \underline{p} = \frac{rs}{rv + \lambda_h(v - s)}.$$

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, despite 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 immediately 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

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

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

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

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 despite 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 $\lambda_\ell = \lambda/2 - \epsilon$ simplifies to:

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

It is immediate that this is continuously differentiable in $\epsilon \in [-\lambda/2, 0)$. 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 $\epsilon \uparrow 0$ and $\epsilon \downarrow 0$ coincide.

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

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

Moreover, differentiating this payoff gap with respect to ϵ and taking its limit as $\epsilon \downarrow 0$ gives us

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

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

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

which is exactly the limit obtained from the limit of a breakthrough-salient environment. Moreover, differentiating the breakdown-salient payoff gap with respect to ϵ and taking its limit as $\epsilon \uparrow 0$ 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 $\epsilon = 0$. ■

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

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

the derivative of which with respect to ϵ is

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

for $\epsilon \leqslant 0$. By the continuity of the first derivative of the payoff gap established in Lemma 2.1, there exists $\tilde{\epsilon} > 0$ such that the gap strictly increases in ϵ for $\epsilon \in (0, \tilde{\epsilon})$ as well.

(ii) In a breakthrough-salient environment $(\lambda_h, \lambda_\ell)$, the payoff gap is $U_a(p_a, p_b) - U_b(p_a, p_b) = A - B$ where $A := 1 - \left(p_a e^{-\lambda_h t^*} + (1 - p_a)e^{-\lambda_\ell t^*}\right) e^{-rt^*}$ and $B := (1 - p_a)\frac{\lambda_\ell}{\lambda_\ell + r}\left(1 - e^{-(\lambda_\ell + r)t^*}\right)(1 + K)$, where K is as defined in the proof of Proposition 2.1. First, the term A strictly decreases in ϵ for any p_a, p_b, ϵ because

$$\frac{\partial A}{\partial\epsilon} = -\frac{p_a\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right)^{-\frac{2\epsilon+\lambda+2r}{4\epsilon}}(\lambda + 2r)\log\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right)}{4\epsilon^2 p_b} < 0.$$

The term A approaches 1 as $\epsilon \to 0$ and $1 - (1 - p_a)\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right)^{-r/\lambda} - p_a\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right)^{-1-r/\lambda} > 0$ as $\epsilon \to \lambda/2$. Second, the term B approaches $\frac{\lambda(1-p_a)(\lambda+p_b\lambda+4r)}{(\lambda+2r)^2}$ as $\epsilon \to 0$ and 0 as $\epsilon \to \lambda/2$. Moreover, (i) for any $\epsilon \in (0, \lambda/2)$, there exists $p_a - p_b$ sufficiently small so that B is arbitrarily close to zero because $B \to 0$ as $p_b \uparrow p_a$ for any $\epsilon > 0$, and (ii) B strictly decreases in ϵ at $\epsilon = 0$ because

$$\left.\frac{\partial B}{\partial\epsilon}\right|_{\epsilon=0} = -\frac{8(1 - p_a)r(\lambda p_b + 2r)}{(\lambda + 2r)^3} < 0$$

and for p_b sufficiently close to p_a, its derivative with respect to ϵ at any $\epsilon > 0$ is arbitrarily close to zero. Therefore, for fixed $(p_a - p_b)$ sufficiently small and for a fixed $\delta > 0$, there exists $\bar{\epsilon} > 0$ such that $0 < A - (U_a(p_a, p_b) - U_b(p_a, p_b)) < \delta$. But A strictly decreases in ϵ, hence $U_a(p_a, p_b) - U_b(p_a, p_b)$ strictly decreases in ϵ for $\epsilon > \bar{\epsilon}$ as well. ■

C Proofs for the large market of section 3

C.1 Proofs for section 3.2 (Fixed wages)

C.1.1 Stable matchings

Stable stage-game matchings. Consider a stage game. Let G denote the CDF of the distribution of the expected productivity p_i of worker $i \in [0, \alpha + \beta]$. Hence, $(\alpha + \beta)G(p)$ is the mass of workers with $p_i \leqslant p$. At time 0, p_i is either p_a or p_b, so $G(p)$ equals 0 if $p < p_b$, $\frac{\beta}{\alpha+\beta}$ if $p_b \leqslant p < p_a$, and 1 if $p \geqslant p_a$. As workers are matched to employers, more is learned about their types, therefore G evolves over time. The evolution of G depends, of course, on the learning environment. In a given stage-game matching, let $d(p)$ denote the fraction of workers with expected productivity p who are matched to an employer.

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, whose productivity is p_s. We look at the unit mass of the most productive workers, and let p^* correspond to the least productive worker in this mass. We will discuss two cases, depending on whether p^* is greater or smaller than p_s.

Definition 3. Fix G. Let p^* be the highest productivity p such that

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

Lemma C.1 shows that worker i is matched if $p_i > \max\{p^*, p_s\}$ and unmatched if $p_i < \max\{p^*, p_s\}$. Therefore, we let $p^M := \max\{p^*, p_s\}$ denote this *marginal productivity*.

Lemma C.1 (Most productive workers are matched)**.** *Fix G.*

1. *Suppose that $p^* > p_s$. Then $d(p)$ equals 1 if $p > p^*$, and 0 if $p < p^*$. If there is no atom at p^*, then $d(p^*)$ can take any value in $[0,1]$. If there is an atom at p^*, then $d(p^*)$ is given by:*

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

   where $G(p^*-) := \lim_{p\uparrow p^*} G(p)$.

2. *Suppose that $p^* \leqslant p_s$. Then $d(p)$ equals 1 if $p > p_s$, and 0 if $p < p_s$. Moreover, $d(p_s)$ can take any value in $[0,1]$ subject to:*

$$(1 - G(p_s))\,(\alpha+\beta) + d(p_s)\,(G(p_s) - G(p_s-))\,(\alpha+\beta) \leqslant 1,$$

   where $G(p_s-) := \lim_{p\uparrow p_s} G(p)$.

Proof. 1. We prove the first part in two steps.

(i) If a less productive worker is matched, then a more productive worker must be matched as well. By way of contradiction, suppose that for a given $p_1 < p_2$, 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.

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

2. Suppose that an unmatched worker has expected productivity $p > p_s$. The mass of workers whose p_i is weakly above p is strictly smaller than 1. Hence, there exists an employer who is either matched to a worker with $p_i < p$ or taking the safe arm. This employer can form a blocking pair with the unmatched worker p. ■

Lemma C.1 characterizes the set of stable stage-game matchings for a given G. This set need not be a singleton. Multiplicity arises from the specification of $d(p^M)$, i.e., the fraction of marginal productivity workers that are matched. If $p^M = p^* > p_s$, multiplicity only arises if G has no atom at p^*, as in part 1 of Lemma C.1. If $p^M = p_s \leqslant p^*$, multiplicity arises because employers are indifferent between a marginal productivity worker and the safe arm, as in part 2 of Lemma C.1. Because stable stage-game matchings are specified uniquely up to the matching of marginal productivity workers, we say that the stable stage-game matching is essentially unique.

Dynamic stability of stable stage-game matchings. In the dynamic setting, G evolves endogenously over time due to learning about workers' types. To argue that 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. Fix any history $h_t \in \mathcal{H}_t$. Let G_{h_t} denote the CDF of p_i at history h_t and $\mu^*_{h_t}$ the essentially unique stable stage-game matching, as in Lemma C.1, given G_{h_t}.

Proposition C.1. *In both the pure breakthrough environment and the pure breakdown one, repeating the stable stage-game matching $\mu^*_{h_t}$ at every h_t is dynamically stable.*

Proof. (i) At h_t, a matched employer's flow payoff is at least s because $p^M(G_{h_t}) \geq p_s$. Hence, she does not strictly prefer taking the safe arm over $[t, t + \mathrm{d}t)$ and then reverting back to $\mu^*_{h_{t+\mathrm{d}t}}$ at $t + \mathrm{d}t$.

(ii) If worker i and employer j are not matched to each other under $\mu^*_{h_t}$ but both strictly prefer to be so over $[t, t + \mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$, then it must be that worker i is unmatched under $\mu^*_{h_t}$ because the wage is fixed, hence $p_i \leq p^M(G_{h_t})$. By the stability of the stage-game matching, any employer is guaranteed at least $p^M(G_{h_t})v$. Hence, employer j must not find it strictly preferable to match with i.

(iii) Suppose that worker i is matched at history h_t according to $\mu^*_{h_t}$. Let $p(t)$ be this worker's expected productivity at history h_t. We focus on the case in which $p(t) \in (0, 1)$, 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 allocated a full task over $[t, \tau)$ for 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,

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

Expressing τ in terms of $p(t)$ from

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

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

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

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 $\tau \geqslant t$. So the worker does not strictly prefer to stay unmatched for $[t, t+\mathrm{d}t)$ and then revert to $\mu^*_{h_{t+\mathrm{d}t}}$.

(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:

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

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 for over $[t, \tau)$ conditional on no breakthrough and the higher the probability that this worker generates a breakthrough over $[t, \tau)$. Staying unmatched over $[t, t + \mathrm{d}t)$ and then reverting to $\mu^*|_{h_{t+\mathrm{d}t}}$ only makes $p(\tau)$ higher than its value on path, so it makes both $\tilde{OM}(\tau)$ and $\tilde{Q}(\tau)$ lower than their values on path. This is true 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 $\mu^*_{h_{t+\mathrm{d}t}}$.

■

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.1.2 Breakthrough learning

Once a worker generates a breakthrough, his employer keeps him for the rest of time. To track how many workers have "secured their jobs", we let $m(t) \in [0, 1]$ 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 $p^M = p_a$. 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 $T_b \downarrow 0$.

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.1.3 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 a *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 for 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. ■

C.2 Proofs for section 3.3 (Flexible wages)

C.2.1 Stable matchings

We first characterize the set of stable stage-game matchings for a fixed G and flexible wages. As in the case of fixed wage, we establish that there exists a unique marginal productivity p^M such that worker i is matched if $p_i > p^M$ and unmatched if $p_i < p^M$. Moreover, worker i's wage is a linear function of p_i.

Lemma C.2 (Equal profit across employers and linear wage for workers)**.** *In any stable stage-game matching,*

1. *all employers make the same profit. If some employers take safe arms, then this profit is s;*
2. *if worker i is matched, his wage takes the form of $p_i v + c_1$, where c_1 is a constant.*

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

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

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

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

Based on Lemma C.2, a stable stage-game matching is without loss characterized by $(d(p), w(p))$, where $d(p)$ specifies the fraction of workers with expected productivity $p_i = p$ who are matched and $w(p) = vp + c_1$ is the wage if a worker with expected productivity p is matched. Moreover, given the linearity of the wage, Lemma C.1 continues to apply intact here as well. Hence, employers are matched to the most productive workers, provided that these workers are better than safe arms.

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

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

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

   (1.a) If for any $\varepsilon > 0$,

   $$\int_{p^*-\varepsilon}^{p^*} (1 - d(s))\, dG(s) > 0,$$

   then $c_1 = -vp^*$ so $w(p_i) = (p_i - p^*)v$.

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

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

Proof. We begin by showing that the wage function must be $w(p_i) = v(p_i - p^*)$ in the case of (1.a). The linearity of $w(p_i)$ follows from Lemma C.2. First, the wage $w(p^*)$ cannot be lower than zero because of limited liability. Second, if $w(p^*) > 0$, then the employer that is matched to p^* worker can form a blocking pair with an unmatched worker whose p_i is arbitrarily close to p^*.

Next we show (1.b). If there exists $\varepsilon > 0$ such that

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

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

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

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

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

C.2.2 Dynamic stability

Lemmata C.1 and C.3 showed that for certain G's there exist multiple stable stage-game matchings. Whenever such multiplicity arises, we select a stable stage-game matching that (i) leaves unmatched marginal-productivity workers as much as possible, and (ii) assigns the employer-preferred wage. This multiplicity arises only at finitely many instants of the entire time horizon. Moreover, the selection criterion that we adopt is for ease of exposition only: the propositions below hold even with a different selection.

Definition 4. Fix G. Let $p^M(G)$ denote the marginal productivity. Let

$$\mu^* := (d^*(p|G), w^*(p|G))$$

be a stable stage-game matching that satisfies the following conditions:

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

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

Proposition C.2. *Under either breakthrough or breakdown learning,* μ^* *is dynamically stable.*

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

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

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

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

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

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

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

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

The worker's expected flow earnings at time τ are

$$\max\left\{0, (1 - Q(\tau))\left(p(\tau) - p^M(G(h_\tau))\right)v\right\} = \max\left\{0, p(t)\frac{p(\tau) - p^M(G(h_\tau))}{p(\tau)}v\right\} \tag{9}$$

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

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

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

The worker's expected flow earnings at time τ are

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

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

■

Our next proposition shows that the contrast between breakthrough and breakdown environments in terms of group inequality continues to hold. In particular, flexible wages do not close the earnings gap between group a and b in the breakdown environment.

Proposition C.3. *Given matching μ^*, as $p_b \uparrow p_a$ the average lifetime earnings of a-workers converge to those of b-workers under breakthroughs but not under breakdowns.*

Proof. Let T_b be as defined in appendix C.1.2. Consider first the breakthrough environment. Because $\alpha > 1$, for an initial period $t \in [0, T_b)$, only a-workers are matched. If an a-worker has not achieved a breakthrough by T_b, his probability of having a high type is p_b. In this case, he has the same continuation payoff as a b-worker does. As $p_b \uparrow p_a$, $T_b \to 0$. Hence, an a-worker's earnings advantage vanishes as well.

We now consider the breakdown environment. Equation (9) in the proof of Proposition C.2 established that a worker who has been matched for longer has higher expected flow earnings than a worker who has been matched for a shorter period. Hence, at any t the expected flow earnings of an a-worker are strictly higher than those of a b-worker. Moreover, group delay T_b does not converge to zero as $p_b \uparrow p_a$, hence an a-worker's earnings advantage due to $[0, T_b)$ does not converge to zero either. Hence, the average lifetime earnings of a-workers are strictly higher than those of b-workers. ■

C.2.3 Wage, earnings, and employment gaps under breakdown learning

In this subsection we normalize v to 1 without loss of generality. We let $E_a(\tau)$ (resp., $E_b(\tau)$) denote the average flow earnings of a-workers (resp., b-workers) at any time $\tau \geqslant 0$. To simplify exposition, we assume that (i) $\alpha > 1$, (ii) $\alpha p_a < 1$, and (iii) $\alpha p_a + \beta p_b > 1$. The first two conditions ensure that the delay for group b is positive but finite, i.e., $0 < T_b < \infty$. The third condition ensures that the pool of new workers is not exhausted before all employers identify a high-type worker. That is, there are more high-type workers than employers available. At the end of this section, we discuss the case of $\alpha p_a + \beta p_b \leqslant 1$.

We first solve for the expected flow earnings at time τ of an i-worker who is first matched at time $t \leqslant \tau$. From expression (9), this expected flow earnings are given by

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

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

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

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

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

In order to calculate the average earnings of i-workers, we also need the density over the time at which each i-worker is first matched. From appendix C.1.3, we have the expression for $\underline{m}(t)$, the mass of workers who have been tried until time t:

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

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

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

which simplifies to:

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

The calculation for the average earnings of b-workers is similar. For any $\tau < T_b$, no b-worker is tried, so the average earnings of b-workers are 0. For $\tau \geqslant T_b$, the average earnings are:

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

Hence,

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

At the start of the horizon, there exists an earnings gap between groups because $E_a(\tau) > 0 = E_b(\tau)$ for any $\tau \in (0, T_b]$. Moreover, the earnings gap persists over the entire horizon and it does not disappear even in the long run, as the following proposition shows. This is because even as $\tau \to \infty$, there exist a non-zero mass of b-workers who never get tried.

Proposition C.4 (Persistent earnings gap under breakdowns)**.** *Suppose that $\alpha > 1 > p_a\alpha$ and $p_a(\alpha+\beta) > 1$. In the limit $p_b \uparrow p_a$, there exists $\tilde{T} \in (T_b, \infty)$ such that the wage gap $W_a(\tau) - W_b(\tau)$ is strictly increasing for $\tau < \tilde{T}$, and strictly decreasing for $\tau > \tilde{T}$. The limit $\lim_{\tau\to\infty}(E_a(\tau) - E_b(\tau))$ is strictly positive.*

Proof. The assumption that $\alpha > 1 > p_a\alpha$ ensures that $T_b \in (0, \infty)$. For any $\tau \in [0, T_b)$, the earnings gap $E_a(\tau) - E_b(\tau)$ is simply $E_a(\tau)$, which is strictly increasing in τ.

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

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

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

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

■

If $\alpha < 1$, then $T_b = 0$. If $\alpha p_a > 1$, then $T_b = \infty$. The results for both cases are similar to those in Proposition C.4, so we omit them. If $p_a\alpha + p_b\beta \leqslant 1$ instead, all b-workers will obtain a chance in the long run. Even though for each $\tau \geqslant 0$ there exists a non-zero earnings gap, as $t \to \infty$ the average earnings of the two groups converge.

We next characterize the average wage of a-workers and that of b-workers at each τ. Let $W_a(\tau)$ and $W_b(\tau)$ be the average wage for the two groups. Let $Q(p_i, \tau - t)$ be the probability that no breakdown has occurred up to time τ if the i-worker is first matched at time t:

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

The average wage of a-workers at time τ is:

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

which simplifies to:

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

The average for b-workers at time $\tau \geqslant T_b$ is:

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

which simplifies to

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

Proposition C.5 (Persistent wage gap under breakdowns)**.** *Suppose that $\alpha > 1 > p_a\alpha$ and $p_a(\alpha+\beta) > 1$. In the limit $p_b \uparrow p_a$, there exists $\hat{T} \in [T_b, \infty)$ such that the wage gap $W_a(\tau) - W_b(\tau)$ is strictly increasing for $\tau < \hat{T}$, and strictly decreasing for $\tau > \hat{T}$.*

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

For any $\tau \in [T_b, \infty)$, we apply the change of variables $x = \frac{p_a-1}{\alpha p_a-1}$, $y = \left(\frac{p_a-1}{\alpha p_a-1}\right)^{-\frac{1}{p_a}} e^{\lambda_\ell\tau}$. We can rewrite the wage gap as

$$\frac{y\left(y^{-p_a}-\frac{x}{y(p_a+x-1)-p_a+1}\right)}{y-1}, \tag{10}$$

where $x > 1$ since $0 < p_a < \alpha p_a < 1$ and $y \geqslant 1$ since $\tau \geqslant T_b$. Note also that y is monotone increasing in τ. This wage gap (10) is increasing in y if and only if

$$H(y) := xy^{p_a}\left(y^2(p_a+x-1)-p_a+1\right)+(-(y-1)p_a-1)(y(p_a+x-1)-p_a+1)^2 > 0.$$

We next argue that $H(y)$ is positive if and only if y is small enough.

First, it is readily verified that $H(1) = H'(1) = 0$, $H(\infty) < 0$, and $H^{(4)}(y) < 0$. This shows that $H^{''}(y)$ is concave. It is also readily verified that $H^{''}(\infty) < 0$. There are three cases to consider regarding the shape of $H''(y)$, with the third case being impossible:

(1) If $H^{''}(1) > 0$, then as y increases, $H^{''}(y)$ is first positive and then negative.

(2) If $H^{''}(1) \leqslant 0$ and $H^{'''}(1) \leqslant 0$, then $H^{''}(y)$ is negative for all $y > 1$.

(3) The last case is $H^{''}(1) \leqslant 0$ but $H^{'''}(1) > 0$. We show that this is not possible since it requires that

$$\begin{aligned} 2(p_a+x) &< p_a x^2 + 2, \\ p_a(x+6)x + 4(x-3)x + 6 &< 6p_a, \end{aligned}$$

which cannot hold simultaneously given that $x > 1$ and $p_a \in (0,1)$.

If case (1) holds, then $H(y)$ is first convex then concave. This, together with $H(1) = H'(1) = 0$ and $H(\infty) < 0$, shows that $H(y)$ is first positive and then negative. If case (2) holds, then $H(y)$ is concave for all $y \geqslant 1$. This, together with $H(1) = H'(1) = 0$, shows that $H(y)$ is negative for $y > 1$. ■

Finally, we also characterize the employment gap between groups. Let $P_a(\tau)$ (resp., $P_b(\tau)$) denote the fraction of a-workers (resp., b-workers) that are allocated a task at time τ. We refer to $P_i(\tau)$ as the *employment rate* for group i. The following proposition shows that at any time τ, a-workers have a strictly higher chance of being employed than b-workers. Moreover, the gap $P_a(\tau) - P_b(\tau)$ does not vanish to zero even as $\tau \to \infty$.

Proposition C.6 (Persistent employment gap under breakdowns)**.** *Suppose that $\alpha > 1 > p_a\alpha$ and $p_a(\alpha + \beta) > 1$. In the limit as $p_b \uparrow p_a$, $P_a(\tau) - P_b(\tau)$ is weakly decreasing in τ and*

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

Proof. The employment rate $P_i(\tau)$ equals $\frac{E_i(\tau)}{W_i(\tau)}$. From the equations for $E_i(\tau)$ and $W_i(\tau)$, we calculate $P_i(\tau)$ as $p_b \uparrow p_a$:

$$P_a(\tau) = \begin{cases} \frac{1}{\alpha} & \text{if } \tau \leqslant T_b \\ p_a + \frac{1}{\alpha}\left(e^{-\lambda_\ell \tau}(1-\alpha p_a)\left(\frac{1-p_a}{1-\alpha p_a}\right)^{1/p_a}\right) & \text{otherwise,} \end{cases}$$

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

The employment gap $P_a(\tau) - P_b(\tau)$ is given by

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

It can be readily observed that (i) for $\tau \leqslant T_b$, $P_a(\tau) - P_b(\tau)$ is constant in τ, (ii) for $\tau > T_b$, it strictly decreases in τ, and (iii) as $\tau \to \infty$, $P_a(\tau) - P_b(\tau) \to \frac{p_a(\alpha+\beta)-1}{\beta}$. Because $p_a(\alpha+\beta) > 1$, this limit is strictly greater than 0. ■

C.2.4 Relaxing limited liability

We have shown that a-workers and b-workers fare quite differently under breakdown learning even if wages are flexible. One might conjecture that this result relies on the assumption that wages have to be nonnegative (that is, the minimum wage must equal the payoff from remaining unemployed): if b-workers could offer negative wages, they would do so and "steal" employment opportunities away from a-workers. In this section, we show that relaxing the limited liability assumption does not guarantee that b-workers have similar employment opportunities as a-workers do, because a-workers will also lower their wages and outbid b-workers. As a result, relaxing the limited liability assumption intensifies competition among workers and thus only benefits the employers.

In this section, we assume that there exists a fixed bound $LB > 0$ such that wages have to be at least $-LB$. We will focus on breakdown learning and show that the disparity between the two groups persists when LB is small enough. (For larger LB, we conjecture that b-workers compete all of their surplus away and have a zero expected lifetime payoff.)

We assume that $\alpha > 1$ and $\alpha p_a < 1$, so according to the dynamic matching μ^* in Proposition C.2 there exists a time $0 < T_b < \infty$ such that b-workers are hired starting from T_b. The marginal productivity is p_a for $t \leqslant T_b$. We also assume that $\alpha p_a + \beta p_b > 1$, so there are more high-type workers than tasks. Due to this assumption, the marginal productivity is p_b for $t \in (T_b, \infty)$.

We revise the dynamic matching μ^* in Proposition C.2 by lowering the wage for a matched worker i at time t from $(p_i - p^M(G(h_t)))v$ to $(p_i - p^M(G(h_t)))v - LB$. Hence, at any time t, the marginal worker's wage is $-LB < 0$. This revised wage function captures the idea that workers benefit from the opportunities to be learnt, so they compete against each other by lowering the wage until the marginal worker's wage drops to the bound $-LB$. This is the only change we made to μ^*. In particular, at any time t, all employers originally had the same flow profit by Lemma C.2. Their flow profit now increases by LB, so all employers continue to have the same flow profit.

We let $\mu^*(LB)$ denote this revised dynamic matching. We next show that if LB is small enough, $\mu^*(LB)$ is dynamically stable.

Proposition C.7. *Assume that $\alpha > 1$, $\alpha p_a < 1$, and $\alpha p_a + \beta p_b > 1$. Under breakdown learning, $\mu^*(LB)$ is dynamically stable for any*

$$
LB < \frac{v(\lambda((2-p_b)p_b - p_a) + r(p_b - p_a))}{\lambda p_b + r}.
$$

In the limit of $p_a \downarrow p_b$, this condition reduces to:

$$LB < \frac{\lambda(1-p_b)p_b v}{\lambda p_b + r},$$

which is equivalent to the condition that a b-worker's continuation payoff at time 0 is strictly positive.

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

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

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

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

If i is not matched, this means that $p_i \leqslant p^M(G(h_t))$. On the other hand, worker i's wage is at least $-LB$, so employer j's flow payoff from being matched to worker i is at most $p_i v + LB$. But employer j's flow payoff on path is at least $p^M(G(h_t))v + LB$. So employer j will not find it strictly profitable to be matched to i.

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

Pick any $\tau > t$. Let $Q(\tau)$ denote the probability that this worker has generated a breakdown in $[t, \tau)$, and $p(\tau)$ denote the worker's expected productivity at time τ conditional on no breakdown in $[t, \tau)$. By Bayes rule,

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

The worker's expected flow earnings at time τ are

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

which is strictly increasing in $p(\tau)$. Staying unmatched over $[t, t+\mathrm{d}t)$ and then reverting to $\mu^*(LB)_{h_{t+\mathrm{d}t}}$ only makes $p(\tau)$ lower than its value on path. Hence, the worker's expected flow earnings at time $\tau \geqslant t+\mathrm{d}t$ is higher on path than if he is unmatched over $[t, t+\mathrm{d}t)$. However, the worker's flow earnings over $[t, t+\mathrm{d}t)$ can be negative if he is matched, so they can be lower than his flow earnings if he is unmatched.

For any $\tau \geqslant t+\mathrm{d}t$, we now compare the worker's expected flow earnings at time τ on and off path. Let $p^{\mathrm{on}}(\tau)$ and $p^{\mathrm{off}}(\tau)$ be, respectively, the probabilities of a high type conditional on no breakdown on path and and off path. Then we have:

$$p^{\text{on}}(\tau) = \frac{p(t)}{p(t) + (1 - p(t))e^{-\lambda(\tau - t)}}$$
$$p^{\text{off}}(\tau) = \frac{p(t)}{p(t) + (1 - p(t))e^{-\lambda(\tau - t - \mathrm{d}t)}}.$$

Substituting $p^{\text{on}}(\tau)$ and $p^{\text{off}}(\tau)$ into (11), we obtain the difference between on-path flow earnings and off-path flow earnings at time τ:

$$p(t)\frac{\left(p^{\text{on}}(\tau) - p^M(G(h_\tau))\right)v - LB}{p^{\text{on}}(\tau)} - p(t)\frac{\left(p^{\text{off}}(\tau) - p^M(G(h_\tau))\right)v - LB}{p^{\text{off}}(\tau)} =$$
$$\left(e^{\lambda \mathrm{d}t} - 1\right)(1 - p(t))e^{-\lambda(\tau - t)}(LB + p^M(G(h_\tau))v) \geqslant \left(e^{\lambda \mathrm{d}t} - 1\right)(1 - p(t))e^{-\lambda(\tau - t)}(LB + p_b v), \quad (12)$$

where the inequality follows from the fact that this payoff difference increases in $p^M(G(h_\tau))$ and that $p^M(G(h_\tau)) \geqslant p_b$. We now integrate the right-hand side of (12) and obtain that the difference between on-path and off-path continuation payoffs at time $t + \mathrm{d}t$ is at least:

$$\int_{t+\mathrm{d}t}^{\infty} e^{-r(\tau - t)}\left(e^{\lambda \mathrm{d}t} - 1\right)(1 - p(t))e^{-\lambda(\tau - t)}(LB + p_b v)\,\mathrm{d}\tau$$
$$= \frac{\left(e^{\lambda \mathrm{d}t} - 1\right)e^{-(\lambda + r)\mathrm{d}t}(1 - p(t))(LB + p_b v)}{\lambda + r} = \frac{\lambda(1 - p(t))(LB + p_b v)}{\lambda + r}\mathrm{d}t + o(\mathrm{d}t). \quad (13)$$

The worker's total discounted earnings in $[t, t + \mathrm{d}t)$ if he stays on path and being matched are:

$$\int_t^{t+\mathrm{d}t} e^{-r(\tau - t)}p(t)\frac{\left(p^{\text{on}}(\tau) - p^M(G(h_\tau))\right)v - LB}{p^{\text{on}}(\tau)}\mathrm{d}\tau \geqslant \int_t^{t+\mathrm{d}t} e^{-r(\tau - t)}p(t)\frac{\left(p^{\text{on}}(\tau) - p_a\right)v - LB}{p^{\text{on}}(\tau)}\mathrm{d}\tau$$
$$= \frac{(p(t) - 1)\left(1 - e^{-(\lambda + r)\mathrm{d}t}\right)(LB + p_a v)}{\lambda + r} - \frac{\left(1 - e^{-r\mathrm{d}t}\right)p(t)(LB - (1 - p_a)v)}{r}$$
$$= ((p(t) - p_a)v - LB)\,\mathrm{d}t + o(\mathrm{d}t), \quad (14)$$

where the inequality follows from the fact that (11) decreases in $p^M(G(h_\tau))$ and that $p^M(G(h_\tau)) \leqslant p_a$. If the worker deviates and stays unmatched, his total discounted earnings in $[t, t + \mathrm{d}t)$ are zero.

The worker prefers to be matched than being unmatched over $[t, t + \mathrm{d}t)$ if the sum of (13) and (14) is positive. For small $\mathrm{d}t$, this is satisfied if

$$LB < \frac{v(p(t)(\lambda(1 - p_b) + r) + \lambda(p_b - p_a) - p_a r)}{\lambda p(t) + r}. \quad (15)$$

The right-hand side increases in $p(t)$, which is the worker's expected productivity at t conditional on no breakdown being realized. Since $p(t)$ is at least p_b, the right-hand side is the smallest when $p(t)$ equals p_b. Hence, the condition (15) is satisfied if

$$LB < \frac{v(\lambda((2 - p_b)p_b - p_a) + r(p_b - p_a))}{\lambda p_b + r}.$$

In the limit of $p_a \downarrow p_b$, this condition reduces to:

$$LB < \frac{\lambda(1-p_b)p_b v}{\lambda p_b + r}. \tag{16}$$

Lastly, we calculate a b-worker's lifetime earnings. Suppose that this worker starts being hired at $t \geqslant T_b$. His expected productivity at time $\tau \geqslant t$ conditional on no breakdown is $p(\tau)$:

$$p(\tau) = \frac{p(t)}{p(t) + (1-p(t))e^{-\lambda(\tau - t)}} = \frac{p_b}{p_b + (1-p_b)e^{-\lambda(\tau-t)}}.$$

Substituting this $p(\tau)$, $p(t) = p_b$, and $p^M(G(h_\tau)) = p_b$ into the worker's expected flow earnings (11) at time τ and integrating the flow earnings over all $\tau \geqslant t$, we obtain the worker's expected lifetime earnings:

$$e^{-rt} \int_t^\infty e^{-r(\tau-t)} p_b \frac{(p(\tau) - p_b)\, v - LB}{p(\tau)} \mathrm{d}t = e^{-rt} \frac{\lambda p_b((1-p_b)v - LB) - LBr}{r(\lambda + r)},$$

which is strictly positive if and only if (16) holds.

■

D Proofs and additional results for section 4

D.1 Auxiliary discussion for section 4.2

Proof of Lemma 4.1. We first show the inequality for the breakdown environment. Suppose $q_a > q_b$, and let $\mu_\ell := \lambda_\ell / r$. The expected payoff of each type of each worker is given by

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

From the definition of the benefit of investment (**??**), it follows that if $q_a > q_b$, then

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

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

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

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

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

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

To characterize $S(h,q_b)$ and $S(\ell,q_b)$, let t_1 be the arrival time of a breakthrough for a high-type worker and let t_2 be the arrival time of his competitor's breakthrough when the task is split equally between workers. For a low type, a breakthrough never arrives. In the absence of any breakthroughs, the employer experiments with the workers until the belief hits $\underline{p}$. The length of this experimentation period is given by t_s as defined above. The CDFs of t_1 and t_2 for $t_1, t_2 \leqslant t_s$ are:[31]

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

with corresponding density functions f_1 and f_2 respectively. Therefore,

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

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

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

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

if $q_a > q_b$, and

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

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

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

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

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

Hence, B_a is continuously differentiable at $q_a=q_b$ as well.[32]

■

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

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

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

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

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

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

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

2. If $c_b < c < c_a$, a low-type worker a is willing to invest but a low-type worker b is not. If the low-type worker a deviates to no investment, the ratio of low-type worker b's payoff to low-type worker a's payoff is κ. By investing worker a obtains a strictly higher payoff. Therefore, the payoff ratio must be strictly lower when the low-type worker a invests.

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

[32]For detailed calculations, see the online supplement at http://yingniguo.com/wp-content/uploads/2020/06/differentiability.pdf.

$\blacksquare$

Proposition D.1 (Investment polarization under breakdown learning). *Fixing all model parameters except for λ_h and λ_ℓ, there exists $\bar{\lambda} > 0$ such that for any $\lambda_h, \lambda_\ell \geqslant \bar{\lambda}$ and in any pair of equilibria, one from each environment, the worker favored post-investment invests strictly more in the breakdown environment than in the breakthrough one and the worker discriminated against post-investment invests strictly less.*

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

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

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

under breakthrough learning, where we use the fact that $\underline{p} \to 0$ as $\lambda_h \to \infty$. The function $\bar{B}_i(q_i, q_{-i})$ increases in q_i, and decreases in q_{-i}. Since q_i is bounded above by $p_a + (1-p_a)\pi$ and q_{-i} is bounded below by p_b, $\bar{B}_i(q_i, q_{-i})$ is bounded from above by

$$\bar{B}_i(p_a + (1-p_a)\pi, p_b) = \frac{(p_a + (1-p_a)\pi)((p_b-2)p_b + 2) - {p_b}^2}{2(p_a + (1-p_a)\pi)(1-p_b)} < 1.$$

By continuity of worker i's benefit from investment in λ_ℓ, λ_h, when λ_ℓ, λ_h are sufficiently large, the worker favored post-investment invests more under breakdown learning than under breakthrough learning.

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

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

under breakthrough learning. Here, the inequality follows from $0 < q_{-i} < 1$. Given that the favored worker i invests more under breakdown than under breakthrough learning, q_i is higher under breakdown learning as well. Hence, the benefit from investment for the worker who is discriminated against is higher under breakthrough learning than under breakdown learning when λ_h, λ_ℓ are large enough. $\blacksquare$

D.2 Proofs for section 4.3

Proof of Proposition 4.3. Let $U_i(p_a, p_b)$ be worker i's payoff given the belief pair (p_a, p_b). For any $p_a > p_b$, the employer first uses worker a for a period of length t^*. If no breakthrough occurs in $[0, t^*)$, the employer's belief toward worker a drops to p_b. Let $f(s)$ for $s \in [0, t^*)$ be the density of the random arrival time of the first breakthrough from worker a. We let $p_a(s)$ be the belief that $\theta_a = h$ if there is no breakthrough up to time s, and let $j(p_a(s))$ be the belief that $\theta_a = h$ right after the first breakthrough at time s. Worker a's payoff is given by

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

Worker b's payoff is given by

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

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

Proof of Proposition 4.4. Let $U_i(q_a,q_b)$ be worker i's payoff given the belief pair (q_a,q_b). We let $p_a(s)$ be the belief toward worker a if there is no breakdown up to time s, and let $j(p_a(s))$ be the belief toward him right after the first breakdown at time s.

Given that $p_a > p_b$, the employer begins with worker a, and uses worker a exclusively if no breakdown occurs. We let $f(s)=p_a\lambda_h e^{-\lambda_h s}+(1-p_a)\lambda_\ell e^{-\lambda_\ell s}$ be the density of the arrival time $s\in[0,\infty)$ of the first breakdown from worker a. We can write worker a's payoff as follows:

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

We can write worker b's payoff as follows:

$$\int_0^{\infty} f(s)e^{-rs}U_b\left(j(p_a(s)),p_b\right)\,\mathrm{d}s.$$

The payoff difference between a and b is:

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

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

$$\int_0^{\infty} f(s)\left(1-2e^{-rs}\right)\,\mathrm{d}s.$$

This term is greater than 0 if and only if $r^2-(1-2p_a)r(\lambda_\ell-\lambda_h)-\lambda_h\lambda_\ell>0$. ■