Early-Career Discrimination: Spiraling or Self-Correcting?

Arjada Bardhi Yingni Guo Bruno Strulovici*

October 17, 2020

Abstract

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

JEL: 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 documented as relying on observable characteristics—such as a worker's gender or race—as statistical proxies for the worker's productivity.[1] Such

*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, 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 the SaMMF Workshop, NU Lunch Seminar, ASU Economic Theory Conference, Bonn-Mannheim Junior Theory Workshop, BFI Economic Theory Conference, and Duke Theory Lunch Workshop 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).

early-career statistical discrimination determines who makes the first cut when opportunities are scarce. Workers from social groups that are expected to be more productive may be systematically prioritized even when 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? One conjecture is that social groups of comparable productivity obtain comparable lifetime earnings: over time, employers learn about workers' productivity after observing their on-the-job performance and reallocate opportunities accordingly. An opposite conjecture suggests that comparable groups fare drastically differently: early opportunities to perform are pivotal in a worker's career progression, and hence, workers favored early on fare substantially better.

This paper shows that the right conjecture depends crucially on how employers learn about workers' productivity. In environments that track 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' 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 when workers can invest in their productivity and, perhaps counterintuitively, even when wages are flexible. Workers' being able to invest in their productivity magnifies the difference in the post-investment productivity of the two groups in learning environments that track failures, but not in environments that track successes. It also creates a *tradeoff between efficiency and equality:* learning environments that alleviate the impact of early-career discrimination also lead to lower equilibrium productivity of employed workers. When wages are flexible—a possibility that we formalize in a dynamic two-sided matching model—comparable groups face very different wage paths in environments that track failures.

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] Sar-

[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 in performing these tasks than do workers who are either employed in other tasks or unemployed, and (iii') groups need not have comparable productivity distributions. Our focus on groups with comparable productivity distributions provides a particularly stark illustration of the

sons (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 the fact that male and female surgeons have comparable abilities: although female surgeons have a lower average ability in her sample, the difference is very small.[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 contributing 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. Another contributing factor is the high cost of setting up and maintaining certain job positions. Task scarcity could also be due to job specialization: only a few workers are needed in highly specialized positions.

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

The employer learns about a worker's productivity from observing the worker's performance. We contrast two learning environments: *breakthrough* and *breakdown*. 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.[4]

The employer's learning environment can be viewed as an intrinsic feature of the job

role played by the learning environment in the dynamics of statistical discrimination.

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

[4]Section 7.1 extends the results to the case in which low-productivity workers also generate breakthroughs but at a lower frequency than high-productivity workers. Similarly, our results remain qualitatively unchanged when high-productivity workers also generate breakdowns but at a lower frequency than low-productivity workers.

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. (Figure 4 in appendix A.1 compares the performance distributions of star jobs and guardian jobs.[5]) Following Jacobs (1981) and Baron and Kreps (1999), scientific researchers and high-stakes salespeople are examples of star jobs, while routine surgeons, airline pilots, and prison guards are examples of 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 the "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 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 his 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 the task. As a result, worker b's expected lifetime earnings are only a fraction of worker a's. 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 becomes arbitrarily fast*. It can explain why societies struggle to eliminate inequality in labor markets.

The contrast between the two employer learning environments is even sharper when workers can invest in their productivity before entering the labor market. The analysis with

[5]Similarly to the performance distribution of star jobs in figure 4, O'Boyle and Aguinis (2012) and Aguinis and Bradley (2015) show that in occupations centered around star performance, such as researchers, 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).

[6]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.

investment is complicated by the fact that multiple equilibria exist in each environment. To address such multiplicity, we compare the lowest payoff inequality attained across all equilibria in each environment. In every equilibrium 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 their productivity in comparable amounts and obtain comparable lifetime earnings. Hence, the self-correcting property of the breakthrough environment persists with investment opportunities.

When learning is sufficiently fast, the breakdown environment features greater polarization in investment incentives than does the breakthrough one. Across all equilibria, the worker who is favored post-investment invests strictly more in the breakdown environment than in the breakthrough one, whereas the worker who is discriminated against post-investment invests strictly less. If, in addition, investment is sufficiently effective, *the employer prefers the breakdown environment*, since it induces almost sure investment by the worker favored post-investment. This result points to a novel tradeoff between efficiency for the employer and equality between the workers.

We further explore this contrast in a large market with many workers from each group and many employers. We show that 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, Flanagan (1978) adapts an argument formulated by Becker (1957) to suggest 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.*

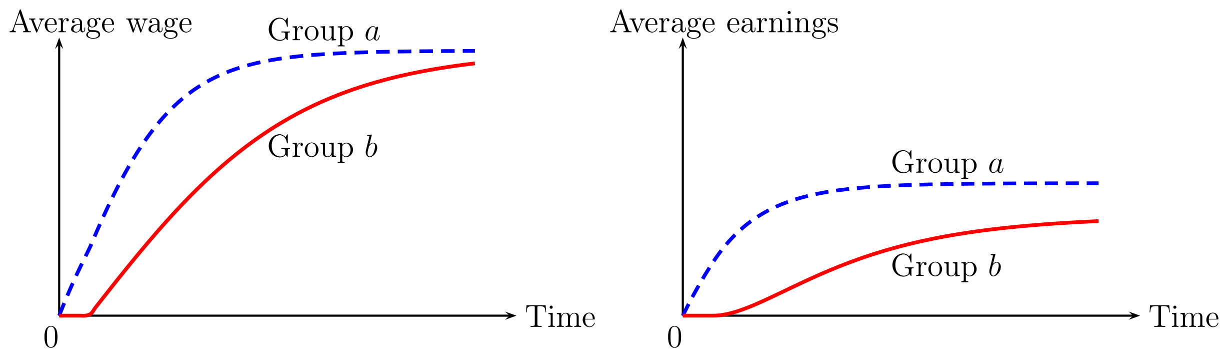

Figure 1: Average wage/earnings under breakdowns for groups of comparable productivity

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 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

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

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.

Early-career discrimination due to prejudice. 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 7.2. 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 literature on statistical discrimination, the theoretical contributions of which are surveyed by Fang and Moro (2011). Phelps (1972) and the literature that followed it (e.g., Aigner and Cain (1977), Cornell and Welch (1996), and Fershtman and Pavan (2020)) assume that there is a significant, exogenous difference between social groups; in these models, inequality between groups arises due to this difference. In contrast, Arrow (1973) and the subsequent literature (e.g., Foster and Vohra (1992), Coate and Loury (1993), Moro and Norman (2004), and Gu and Norman (2020)) assume no exogenous difference between groups; in these models, 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 Phelps (1972) and its subsequent literature, inequality across groups disappears as the difference between groups vanishes, whereas our model highlights the fact that a vanishingly small difference can snowball into

[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.

a large payoff gap. Second, in contrast to Arrow (1973) and its subsequent literature, the across-group inequality that we uncover is not due to the existence of multiple equilibria. Third, most papers in both of these 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. 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.

The paper contributes to the literature on employer learning (e.g., Farber and Gibbons (1996), Altonji and Pierret (2001)). The employer's learning environment can be interpreted as an intrinsic feature of an occupation. In this respect, our work is related to Altonji (2005), Lange (2007), Antonovics and Golan (2012), and Mansour (2012). Whereas the models in these papers 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 this distinction.

Our analysis leverages the tractability of Poisson bandits, which have been used widely in strategic experimentation models (e.g., Keller, Rady and Cripps (2005), Keller and Rady (2010), Strulovici (2010), Keller and Rady (2015)).[10] In our setting, the employer is the bandit operator and the workers are the bandit arms. Because we explore workers' incentives to invest in productivity, the quality of the bandit arms is endogenously determined. In modeling bandit arms as strategic players, our paper is related to Bergemann and Valimaki (1996), Felli and Harris (1996), and Deb, Mitchell and Pai (2019). Unlike our analysis of investment in productivity, these models assume that the quality of the bandit arms is exogenously given.[11] One exception is Ghosh and Hummel (2013), in which the quality of the arms is endogenous.

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

[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]Both Felli and Harris (1996) and this paper use the framework of multi-armed bandits to model labor market learning.

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.

Structure of the paper. Section 2 presents the baseline model. Section 3 identifies the contrast between breakthrough and breakdown environments. Section 4 analyzes the role of workers' investment in productivity. Section 5 generalizes the model to many workers from each group and many employers. Section 6 shows that allowing for flexible wages does not overturn the contrast drawn in section 3. Section 7 establishes the robustness of our results to more general learning environments and to misspecified beliefs. Section 8 concludes.

2 Baseline model

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

Before time $t = 0$, 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 workers' types. We interchangeably refer to p_i as the *prior belief* for worker i or as worker i's *expected productivity* at time 0. We assume that worker a is ex ante more productive: $p_a > p_b$. Our focus is on groups with comparable expected productivity, i.e., when p_b is close to p_a.

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

A worker obtains a flow payoff $w > 0$ whenever he is assigned the task. Otherwise, his flow payoff is zero. The parameter w can be interpreted as the fixed wage for a worker who performs the task. Without loss of generality, we normalize w to one. Section 6 provides a general analysis of the case in which workers' wages are flexible.

The employer obtains a flow payoff $v > 0$ if she allocates the task to a high-type worker, and a flow payoff normalized to zero if she allocates the task to a low-type worker. These payoffs are the employer's net payoffs after wage w 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]

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

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

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

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

A breakthrough perfectly reveals a high-type worker and a breakdown perfectly reveals a low-type one. In section 7.1, we extend our analysis to inconclusive breakthrough environments ($\lambda_h > \lambda_\ell > 0$) and inconclusive breakdown environments ($\lambda_\ell > \lambda_h > 0$).[13]

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

We let $\underline{p}$ denote the belief below which the employer switches to the safe arm. This belief threshold, derived in in appendix A.2, is given by

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

The threshold $\underline{p}$ is lower than the myopic threshold s/v in both environments due to the value of learning for future allocation decisions. 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.

[12] We scale players' lifetime payoffs by a factor of r, as per standard practice in experimentation models (e.g., 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.

[13] In our formulation, the employer learns through observing signals rather than her payoffs. This is equivalent to an alternative formulation in which the employer learns through observable payoffs. In this alternative formulation: (i) 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, (ii) 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.

3 Benchmark Comparison

This section establishes a stark contrast between the two learning environments. We analyze workers' expected lifetime payoffs given the employer's optimal task allocation behavior. Our focus is on groups with comparable productivity distributions: we compare workers' payoffs when the expected productivity of one group is arbitrarily close to that of the other and study how this comparison depends on whether the signal takes the form of a breakthrough (section 3.1) or a breakdown (section 3.2).

3.1 Self-correction under breakthrough learning

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

At each instant, the employer allocates the task to the worker with the higher expected productivity.[14] 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. Here, t^* measures how long it takes for the belief about worker a's type to drift down from p_a to p_b in the absence of a breakthrough. If worker a generates a breakthrough before t^*, the employer allocates the task 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.

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

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

Proof. The employer initially allocates the task exclusively to worker a. If worker a produces no breakthrough, this initial allocation lasts until the employer's belief that a has a

[14]This is true in both learning environments 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.

high type decreases to p_b, which happens at time t^*, where t^* is defined by

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

After t^*, the employer splits the task equally between the two workers until either a breakthrough occurs or the belief about the workers' types hits $\underline{p}$. If a breakthrough occurs, the employer thereafter allocates the task only to the worker who generated the breakthrough.

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

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

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

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

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

3.2 Spiraling under breakdown learning

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

Proposition 3.2 establishes a *spiraling property* of breakdown learning: even if p_b is just slightly less than p_a, worker a 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 breakthrough environment, in which worker a loses his preferential status if he fails to generate a breakthrough within a given time window.

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

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

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

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

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:

$$(1-p_a)\frac{\lambda_\ell}{\lambda_\ell+r}\left(p_b+(1-p_b)\frac{r}{\lambda_\ell+r}\right). \tag{4}$$

As $p_b\uparrow p_a$, worker b's expected payoff converges to

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

■

For spiraling to arise, it is crucial that the delay faced by worker b not depend on how close p_b is to p_a. The payoff ratio (2) has two components: (i) $(1-p_a)$ reflects the fact that worker b obtains a chance only if worker a has a low type, and (ii) $\lambda_\ell/(\lambda_\ell+r)$ reflects the expected time it takes for worker a's low type to be revealed. Moreover, even as learning becomes instantaneous—i.e., as $\lambda_\ell\to\infty$—this payoff ratio approaches $(1-p_a)$ rather than one. Being the second hire is detrimental to worker b even when the employer learns very fast: worker b never obtains a chance if worker a has a high type.

Since groups have similar productivity distributions, even if the employer were blind to which group each worker belonged to and treated the workers equally, her payoff would not be much lower than when she observes who belongs to each group. In the limit as $p_b\uparrow p_a$, her payoff would be the same in both cases. Therefore, making it more difficult for the employer to observe who belongs to each group would equalize workers' payoffs without making the employer worse off.[15]

4 Investment in productivity

In this section workers can invest in their productivity before entering the labor market. We explore the equilibrium implications of this investment opportunity for workers' prospects. These implications are unclear a priori: access to investment might level the playing field or it might amplify the expected productivity gap between workers. Yet again, it turns out that the learning environment plays a key role.

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

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

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

4.1 Investment equilibria

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

$$B_i(q_a, q_b) := \pi \left(U_i(h; q_a, q_b) - U_i(\ell; q_a, q_b) \right), \tag{5}$$

[15] 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.

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

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

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

Both learning environments have the following feature: 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.[16] This is because worker i would be the first to be allocated the task: by investing, he might avoid a breakdown in the near future or increase the chance of a breakthrough within the grace period allotted exclusively to him.

Lemma 4.1 (The worker favored post-investment has higher benefit of investment)**.** *In both learning environments, if $q_i > q_{-i}$, then $B_i(q_a, q_b) > B_{-i}(q_a, q_b)$. For each i, $B_i(q_a, q_b)$ is continuously differentiable in the breakthrough environment, but it is not continuous at $q_a = q_b$ in the breakdown environment.*

Proof of Lemma 4.1. 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 (5), it follows that if $q_a > q_b$, then

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

[16] 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 $q_a = q_b$.

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 $q_i = q_{-i}$.

The proof for the breakthrough environment, to be found in appendix B, follows similar steps but is algebraically more involved. ∎

The worker who is favored by the employer after the investment stage has a stronger incentive to invest. This, 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, when worker a's expected productivity advantage prior to investment is sufficiently small, there exist equilibria in which worker b invests more than worker a and becomes favored post-investment. Therefore, investment can reverse the initial ranking of groups. To address this multiplicity, in each environment we characterize the lowest payoff inequality attained across all equilibria as $p_b \uparrow p_a$. The next subsection shows that this lowest payoff inequality continues to be zero in the breakthrough environment. In the breakdown environment, however, it is even greater than the payoff inequality in the no-investment benchmark.

When $p_a \neq p_b$, there cannot exist an equilibrium in which $q_a = q_b$. If such an equilibrium existed, investment would provide the same benefit to both workers. Workers would therefore use the same investment strategy. However, since they have unequal probabilities of having a high type pre-investment (i.e., $p_b < p_a$), worker b would need to invest more in order to attain $q_a = q_b$, which means that the workers' investment strategies would have to be different.

4.2 An even starker contrast with investment

4.2.1 Self-correction under breakthrough learning

Our next proposition formalizes the notion that the self-correcting property of the breakthrough environment continues to hold with investment. It does so by establishing that there always exists an equilibrium in which q_b converges to q_a as $p_b \uparrow p_a$. In this equilibrium the workers' benefits from investment get arbitrarily close, and so do their investment thresholds. Hence, the payoff gap between comparable workers becomes vanishingly small.

This equilibrium could either preserve the prior ranking of the workers—with worker a being favored post-investment—or reverse it.

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

Genericity here is meant as follows: fixing all parameters of the model except for (p_a, π), the set of values of $(p_a, \pi) \in \big(\underline{p}, 1\big) \times (0, 1)$ for which the proposition does not hold has measure zero.[17] The proof of the proposition builds on two observations. First, when the pre-investment probabilities are the same (i.e., $p_a = p_b$), there always exists a symmetric equilibrium in which the workers use the same cost threshold and thus have the same post-investment probability of having a high type. Second, under breakthrough learning the benefit from investment is continuously differentiable in (q_a, q_b). We invoke 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. The proof is standard, hence relegated to online appendix D. The other proofs for this section are in appendix B.

4.2.2 Exacerbated spiraling under breakdown learning

In contrast, access to investment not only fails to tame the propensity of breakdown learning to magnify small prior differences, but makes it worse. Across *all* equilibria, the expected payoffs of ex ante comparable workers are even further apart than in the no-investment benchmark of section 3.2.

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

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

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

Unlike the case in the breakthrough environment, the benefit from investment in the breakdown environment is discontinuous in (q_a, q_b), as shown in Lemma 4.1. This difference

[17] When 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.

explains why a proof similar to that for Proposition 4.1 does not work here. Nonetheless, the benefit function takes a simple form—as already previewed in the proof of Lemma 4.1—hence, characterizing the set of equilibria is straightforward. As $p_b \uparrow p_a$, there exist only two equilibria: worker a is favored post-investment in one equilibrium and worker b is favored post-investment in the other. These two equilibria are the same modulo the workers' identities.

As we saw in Proposition 3.2, the payoff ratio across workers in the no-investment benchmark is pinned down by p_a, the probability that worker a has a high type without investment. 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. We show that 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 post-investment—after factoring in the investment cost—is lower than the ratio in the no-investment benchmark. Hence, inequality between workers increases due to the investment opportunity.

4.3 Polarization in investment behavior under breakdown learning

Our discussion has so far focused on comparing learning environments in terms of workers' payoffs. We now turn to differences in investment behavior. In a nutshell, when learning is sufficiently fast, the worker favored post-investment invests strictly more under breakdowns than under breakthroughs, whereas the worker discriminated against post-investment invests strictly less under breakdowns. Therefore, the breakdown environment is marked by greater polarization in workers' investment behavior. Notably, this comparison holds across all equilibria and does not hinge on the arrival rates being equal across environments. All that is needed is that the arrival rates be sufficiently high in both environments.

Proposition 4.3 (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.*

When learning is sufficiently fast, investment incentives are seemingly similar across the two environments. The worker favored post-investment is the first to be allocated the task and information about his type arrives quickly, so his incentives to invest are quite strong. For the worker discriminated against post-investment, investment incentives are less clear. On the one hand, a more strongly favored worker depresses the investment incentives of the worker discriminated against. On the other, due to fast learning, the worker who is discriminated against might get a chance earlier if he invests. Despite these seeming similarities across environments, Proposition 4.3 identifies a key difference between them.

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

The worker who is discriminated against post-investment, on the other hand, has a weaker incentive to invest under breakdowns than under breakthroughs when learning is sufficiently fast. This is due to two forces that reinforce each other. First, because of spiraling, the breakdown environment already makes the second worker much less likely to be hired than in the breakthrough environment. Second, under breakdowns the worker who is discriminated against faces a competitor who invests strictly more—as explained above—which further lowers the chance that this worker will get a shot from the employer.

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

Corollary 4.2 (Employer prefers breakdown learning)**.** *Fixing all model parameters except for λ_h, λ_ℓ, and π, there exists $\bar{\lambda} > 0$ and $\bar{\pi} \in (0,1)$ such that for any $\lambda_h, \lambda_\ell \geqslant \bar{\lambda}$ and $\pi > \bar{\pi}$, the employer's payoff is strictly higher under breakdowns than under breakthroughs.*

Note that in the no-investment benchmark, the employer's payoff is the same across the two environments as λ_h, λ_ℓ become very large: her payoff equals the probability that at least one worker has a high type. This shows that the employer's strict preference for the breakdown environment in the presence of investment is due to the difference in investment incentives across the two environments.

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

5 Many employers and workers: The role of task scarcity

Our baseline model focused on a single employer choosing between two workers from distinct groups. We now consider a market with many employers and many workers from each of the two groups (still modeling only two groups for simplicity).[18] Comparable groups continue to have comparable payoffs under breakthroughs, but markedly different payoffs under breakdowns. Moreover, the scarcer tasks become relative to workers, the greater the inequality between groups in the breakdown environment.

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

Moreover, we assume that $\alpha + \beta > 1$, so tasks are scarce and some workers are not initially hired. Such task scarcity is both necessary and sufficient for payoff inequality to emerge in the breakdown environment. For ease of exposition, we further assume $\alpha > 1$ for the rest of this discussion. The formal framework for this section is presented in appendix C.1 and the rest of the analysis is in online appendix E.

Diverse hiring under breakthrough learning. Mirroring the analysis in section 3.1, the dynamic allocation of tasks goes through two phases. In the first phase, all a-workers take turns to perform tasks. If an a-worker generates a breakthrough, he "secures his job" with his current 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

[18]The results of this section can be readily extended to more than two groups.

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.

Narrow hiring under breakdown learning. Breakdown learning, in contrast, leads to sluggishness in trying 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 reported in section 3.2 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.

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 5.1 states 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 5.1 (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.[19]

6 Flexible wages

This section investigates whether flexible wages can restore earnings equality in the breakdown environment.[20] We show that, as long as workers' wages are nonnegative, both the self-correcting property of breakthroughs and the spiraling property of breakdowns still hold. In particular, wage flexibility is insufficient to prevent spiraling.

We incorporate flexible wages into the dynamic, large market from section 5. In this richer setting, a stage-game outcome specifies not only how workers are matched to employers but also a wage for each matched pair. We call this a *stage-game matching*. Using the solution concept of Shapley and Shubik (1971), we first characterize the stable stage-game matching, which is essentially unique. The dynamic counterpart of a stage-game matching—a *dynamic matching*—specifies, after each history, how workers are matched to employers and a wage for each matched pair. Adopting the solution concept of Ali and Liu (2020), we show that *prescribing the stable stage-game matching after each history is dynamically stable*: no worker-employer pair and no single player has a profitable one-shot deviation after any history. The formal framework is presented in appendix C.2. The rest of the analysis is in online appendix F.

[19]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.

[20]Section 6 assumes that workers do not know their types at time 0: they share the same prior belief as the employers. All players learn symmetrically about a worker's type. 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).

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 $p^M(t)v$.[21]

Dynamic stability of the stable stage-game matching. Why is prescribing the stable stage-game matching after each history 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 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 $\max\left\{0, (p_t - p^M(t))v\right\}$, as figure 2 shows.[22] 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.

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

[21]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).

[22]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 2 would be reinterpreted as *net* flow earnings.

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 2. In the first period, a-workers 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.

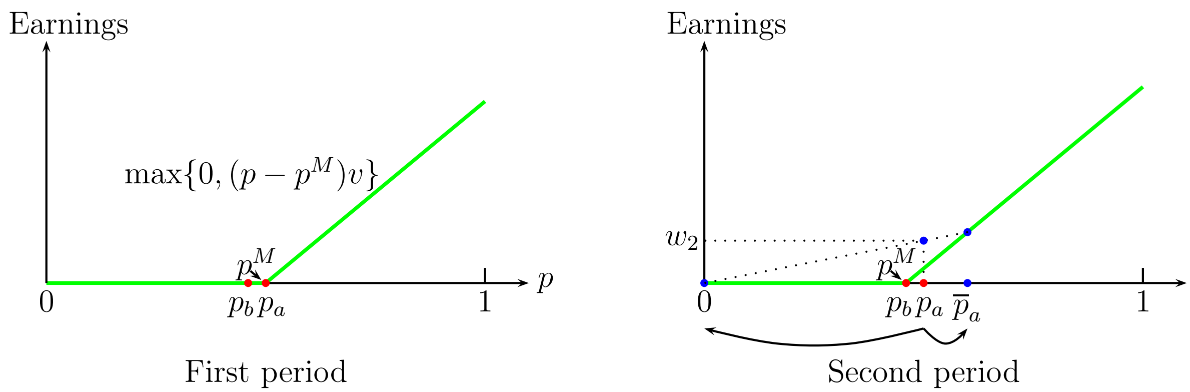

Figure 2: A two-period example with $\alpha = \beta = 1$

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. Online appendix F.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 F.5). 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 F.4). See figure 3 below for an illustration.[23]

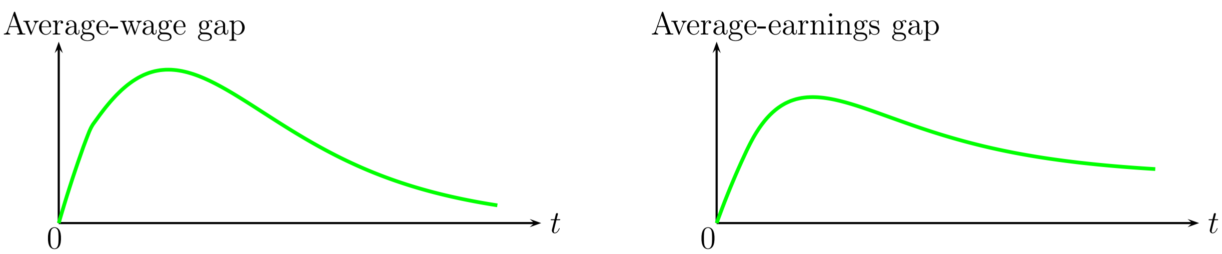

Figure 3: Average-wage gap and average-earnings gap as $p_b \uparrow p_a$

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 F.3). 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 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 $t \to \infty$.

7 Alternative learning and belief specifications

7.1 Inconclusive learning environments

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

[23] Figure 1 in the introduction and figure 3 assume the same parameter values: $\alpha = 5/4$, $\beta = 1$, $p_a = 1/2$, $\lambda_\ell = 1$, and $r = 1$.

the signal suggests a high type (i.e., $\lambda_h > \lambda_\ell$) and an *inconclusive breakdown environment* otherwise (i.e., $\lambda_h < \lambda_\ell$). If $\lambda_h = \lambda_\ell$, signals are uninformative.

The self-correcting property extends to inconclusive breakthrough environments, as established in Proposition G.1 in online appendix G. 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 $p_b \uparrow p_a$, 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.

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.[24] Proposition G.2 presents the details.

7.2 Misspecified prior belief

Suppose that the two workers have the same probability p_{true} of having a high type, but the employer believes that worker b has a lower probability $p_{\text{mis}} < p_{\text{true}}$.[25] The spiraling property of the breakdown environment continues to hold, in the sense that even a very slight misspecification grants a large payoff disadvantage to worker b. Worker a is still hired first based on the employer's misspecified belief and the workers' payoffs are still given by expressions (3) and (4) (with p_a and p_b being both replaced by p_{true}).

The self-correcting property of the breakthrough environment continues to hold as well, in the sense that a slight misspecification will not have large payoff consequences for workers. Duration t^*, which is analogous to the grace period in (1), 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

[24] 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.

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

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

We let $\hat{U}_a(p_{\mathrm{mis}}, p_{\mathrm{true}})$ and $\hat{U}_b(p_{\mathrm{mis}}, p_{\mathrm{true}})$ 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 $\hat{U}_a(p_{\mathrm{mis}}, p_{\mathrm{true}})$ is lower than $\hat{U}_b(p_{\mathrm{mis}}, p_{\mathrm{true}})$. Crucially, as p_{mis} converges to p_{true}, the two payoffs get arbitrarily close. To extend the proof of Proposition 3.1 to the misspecified-prior case, we only need to replace t^* with the new definition and replace $U_i(p_b, p_b)$ with $\hat{U}_i(p_{\mathrm{mis}}, p_{\mathrm{true}})$ for 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.

8 Concluding remarks

This paper studies the consequences of different employer learning environments for social groups of comparable expected productivity. Whether the learning environment is closer to a breakdown environment or a breakthrough one has important implications for whether discrimination persists in the long run. Lange (2007) observed that "how economically relevant statistical discrimination is depends on how fast employers learn about workers' productive types." Our analysis provides an additional perspective: what matters for statistical discrimination is not only the speed of employer learning, but also the 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. Corollary 4.2 describes circumstances under which the employer prefers breakdown learning. More generally, is the endogenous choice of the learning environment more likely to lead to breakdown or breakthrough learning? 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.

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.

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 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. 2019. "Our Distrust is Very Expensive."

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.

Ghosh, Arpita, and Patrick Hummel. 2013. "Learning and Incentives in User-Generated Content: Multi-Armed Bandits with Endogenous Arms."

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

Gu, Jiadong, and Peter Norman. 2020. "A Search Model of Statistical Discrimination."

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

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

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

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

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

Kalman, Peter J., and Kuan-Pin Lin. 1979. "Applications of Thom's Transversality Theory and Brouwer Degree Theory to Economics." *Journal of Mathematical Analysis and Applications*, 67(2): 249–260.

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.

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

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

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

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

Moro, Andrea, and Peter Norman. 2004. "A General Equilibrium Model of Statistical Discrimination." *Journal of Economic Theory*, 114: 1–30.

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.

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

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

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

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

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

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 4 depict the probability density of performance signals for a guardian job and that for 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' type, but they would look similar if the probabilities were conditional on a low type under breakdowns (figure 4a) and conditional on a high type under breakthroughs (figure 4b). 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 5 for examples of such empirical studies.

A.2 Derivation of $\underline{p}$ in section 2

Lemma A.1. *The belief threshold at which the employer switches to the safe arm is given by:*

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

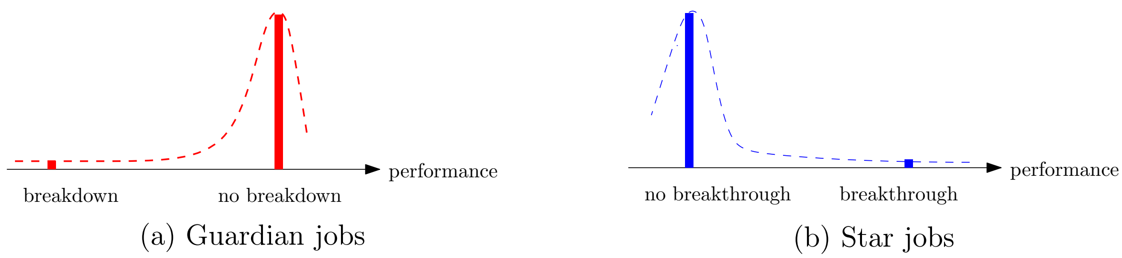

Figure 4: Distribution of performance signals (adapted from Baron and Kreps (1999))

Proof. Consider first $\lambda_h > \lambda_\ell = 0$. Fixing an arbitrary prior belief p and threshold belief $\underline{p} < p$, this corresponds to duration

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

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

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

whereas conditional on the worker having a low type, the employer's payoff is $e^{-rt^*\left(p,\underline{p}\right)} s$. Hence, the expected payoff of the employer simplifies to

$$V_{BT}(p, \underline{p}) := pv + t^*\left(p, \underline{p}\right)^{-(\lambda_h + r)/r}(1 - pv).$$

The smooth pasting condition yields

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

Next, consider $\lambda_\ell > \lambda_h = 0$. If the worker has a high type, the payoff of the employer is v. 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

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

■

B Proofs for section 4

Continuation of the proof of Lemma 4.1. We now show the inequality for the breakthrough environment. Let $q_a > q_b$. The employer uses worker a exclusively for a period of length $t^* = \frac{1}{\lambda_h} \log\left(\frac{q_a(1-q_b)}{(1-q_a)q_b}\right)$ and then splits the task equally among the two workers for a subsequent period of length $t_s := \frac{2}{\lambda_h} \log\left(\frac{q_b(1-\underline{p})}{(1-q_b)\underline{p}}\right)$. 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:[26]

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

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

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

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

if $q_a > q_b$, and

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

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.[27]

■

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

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

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.

■

Proof of Proposition 4.3. Throughout the proof, we set $\pi = 1$ without loss, as π merely scales the benefit from investment $B_i(q_a, q_b)$ and 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. ■

C Framework for sections 5 and 6

C.1 Large market framework of section 5

This section extends our baseline model to a two-sided matching market with a continuum of workers and employers. There is a unit mass of employers, a mass of size $\alpha > 1$ of a-workers, and a mass of size $\beta > 0$ of b-workers. Both employers and workers are long-lived. They 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. An a-worker's type is high with probability p_a, and a b-worker's type is high with probability p_b. There is also a unit mass of identical safe arms available. We assume that $p_a > p_b > \underline{p}$.

At each instant, each employer has one task to allocate and each worker can take up at most one task.

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

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

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

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

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

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

An employer returns to the unmatched pool if she fires the worker that she is currently matched to. We assume that all signals are observed publicly. Therefore, all employers hold the same belief about the type of any worker at any time t. The frictionless matching protocol ensures that at each instance the most productive workers are matched provided that their probabilities of having a high type are higher than $\underline{p}$. As in the baseline model, employers are willing to experiment at a belief that is as low as $\underline{p}$, because (i) an employer reaps the entire benefit from her worker's successful experimentation in the same manner as in the baseline model, and (ii) workers' types are independent.

Task scarcity. When there are more workers than tasks (i.e., $\alpha + \beta > 1$), not all workers are matched immediately at $t = 0$. Such relative scarcity of tasks is both necessary and sufficient for spiraling to arise under breakdown learning. To simplify exposition, we impose the stronger assumption that $\alpha > 1$.[28] This assumption guarantees that only a-workers are matched at $t = 0$. The rest of the analysis is in online appendix E.

C.2 Flexible wage framework of section 6

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

In this section, we use $i \in [0, \alpha + \beta]$ to index a worker. Worker i is from group a if $i \in [0, \alpha]$ and from group b if $i \in (\alpha, \alpha + \beta]$. We use $j \in [0, 1]$ to index an employer. At each instant, each employer has one task to allocate and each worker can take up at most

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

one task. There is also a unit mass of identical safe arms. An employer takes a safe arm whenever not matched to a worker. We assume that no one observes the workers' realized types and that all learning is public.

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

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

Throughout this section, we make the following assumption. As it will become clear in subsection F.1, if a worker's expected productivity (i.e., probability of having a high type) is below p_s, he is not matched to employers in the stable stage-game matching.

Assumption 1. *An employer's flow payoff from an a-worker or a b-worker given the prior beliefs p_a, p_b is higher than that from the safe arm: $p_s < p_b < p_a$.*

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

Solution concept. We first adopt the solution concept in Shapley and Shubik (1971) and define stable stage-game matchings. In appendix F.1 we 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 following (D, W).

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

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

(ii) there exists no blocking pair.

Next we define dynamic stability based on the solution concept of a *stable convention* in Ali and Liu (2020).[29] For a given dynamic matching μ, let $\mu|_h$ denote the continuation matching after some history h.

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

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

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

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

The rest of the analysis is in online appendix F. In particular, appendix F.1 characterizes the set of stable stage-game matchings, whereas appendix F.2 establishes that repeating a stable stage-game matching is dynamically stable in both learning environments. For such dynamically stable matchings, we show that the self-correcting property of breakthrough learning and the spiraling property of breakdown learning continue to hold. Appendix F.3 analyzes the employment, wage, and earnings gaps under this dynamically stable matching in the breakdown environment.

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

Online appendix

Early-Career Discrimination: Spiraling or Self-Correcting?

Arjada Bardhi Yingni Guo Bruno Strulovici

October 17, 2020

D Proof of Proposition 4.1

Proof of Proposition 4.1. A post-investment belief pair (q_a, q_b) and a cost-threshold pair (c_a, c_b) constitute an equilibrium if and only if $\forall i \in \{a, b\}$:

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

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

$$\begin{cases} \dfrac{1}{\pi} B_a\left(q_a, q_b\right) - \dfrac{1}{\pi} F^{-1}\left(\dfrac{q_a - p_a}{(1-p_a)\pi}\right) & = 0 \\ \dfrac{1}{\pi} B_b\left(q_a, q_b\right) - \dfrac{1}{\pi} F^{-1}\left(\dfrac{q_b - p_b}{(1-p_b)\pi}\right) & = 0. \end{cases} \tag{6}$$

Let $g_a(p_a, p_b, q_a, q_b)$ and $g_b(p_a, p_b, q_a, q_b)$ denote respectively the LHS of each equation in (6). Both g_a and g_b are continuously differentiable, because B_a, B_b and F are continuously differentiable and F' is strictly positive.

Existence of symmetric equilibrium. We first show that if workers have the same prior belief, there is a symmetric equilibrium in which they have the same post-investment belief. Let $\hat{p}$ denote the two workers' prior belief and define

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

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.

A symmetric equilibrium exists if there exists $\hat{q} \in [\hat{p}, \hat{p} + (1 - \hat{p})\pi]$ such that $g(\hat{q}, \pi) = 0$, or equivalently,

$$\frac{\pi \left(\mu_h + \frac{\left(\frac{\hat{q}(1-\underline{p})}{(1-\hat{q})\underline{p}}\right)^{-\frac{\mu_h+2}{\mu_h}} ((\mu_h+2)\hat{q}+\underline{p}(\mu_h(\underline{p}-2)-2))}{(1-\underline{p})\underline{p}} \right)}{2(\mu_h+2)} = F^{-1}\left(\frac{\hat{q}-\hat{p}}{\pi(1-\hat{p})}\right). \tag{7}$$

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

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

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

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

$$\left.\frac{\partial g_a}{\partial q_a} + \frac{\partial g_a}{\partial q_b}\right|_{(\hat{p},\hat{p},\hat{q},\hat{q})} \neq 0 \tag{8}$$

$$\left.\frac{\partial g_a}{\partial q_a} - \frac{\partial g_a}{\partial q_b}\right|_{(\hat{p},\hat{p},\hat{q},\hat{q})} \neq 0. \tag{9}$$

Claim (8) holds because

$$\left.\frac{\partial g(q, \pi)}{\partial q}\right|_{q=\hat{q}} < 0.$$

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

Lemma D.1. *There exists a set $\Pi \subset (0, 1)$ of measure one such that $g(q, \pi)$ intersects zero transversally at each intersection point for any $\pi \in \Pi$.*

Proof. First, $g(q, \pi)$ is strictly increasing in π because the term $\frac{1}{\pi}B_i(q, q)$ is independent of π and F^{-1} is strictly increasing in $[0, 1]$. Therefore 0 is a regular value of $g(q, \pi)$. By the Transversality Theorem (Kalman and Lin (1979)), there exists a set $\Pi \in (0, 1)$ of values for π such that $(0, 1) \setminus \Pi$ has measure zero and for any $\pi \in \Pi$, 0 is a regular value of $g(q, \pi)$.

Hence, generically the derivative of $g(q, \pi)$ with respect to q at any intersection point $q = \hat{q}$ such that $g(\hat{q}, \pi) = 0$ is non-zero. ■

Claim (9) holds unless:

$$\frac{\frac{\left(\frac{\hat{q}(1-\underline{p})}{(1-\hat{q})\underline{p}}\right)^{-2/\mu_h}\left((\mu_h+2)\hat{q}^2+\mu_h(2\hat{q}-1)\underline{p}^2-2(\mu_h+1)(2\hat{q}-1)\underline{p}\right)}{(\underline{p}-1)^2}+\frac{2\hat{q}(\mu_h\hat{q}+1)}{1-\hat{q}}}{2(\mu_h+2)\hat{q}^2} = \frac{1}{\pi^2(1-\hat{p})F'\left(F^{-1}\left(\frac{\hat{q}-\hat{p}}{\pi(1-\hat{p})}\right)\right)}. \tag{10}$$

Fix $(F, \underline{p}, \mu_h)$. The following lemma shows that for almost any $(\pi, \hat{p})$ claim (9) holds.

Lemma D.2. *Suppose that F is weakly convex. Then, claim (9) is satisfied in equilibrium for almost all $(\pi, \hat{p})$.*

Proof. The system of equations (7) and (10) is equivalent to:

$$g_1(\hat{p}, \hat{q}, \pi) := \frac{1}{\pi}F^{-1}\left(\frac{\hat{q}-\hat{p}}{(1-\hat{p})\pi}\right) - h_1(\hat{q}) = 0$$
$$g_2(\hat{p}, \hat{q}, \pi) := \frac{1}{\pi(1-\hat{p})F'(\pi h_3(\hat{q}))} - h_2(\hat{q}) = 0,$$

where h_1, h_2, h_3 are functions of $\hat{q}$ only and h_3 is defined from the equilibrium condition (6) as:

$$h_3(\hat{q}) := \frac{1}{\pi}F^{-1}\left(\frac{\hat{q}-\hat{p}}{(1-\hat{p})\pi}\right) = \frac{1}{\pi}B_a(\hat{q}, \hat{q}).$$

Note that g_1 is strictly decreasing in $\hat{p}$ and π, whereas g_2 is strictly increasing in $\hat{p}$ but decreasing in π, by the convexity of F. Therefore, the determinant of the Jacobian matrix of this system with respect to $(\pi, \hat{p})$ is strictly negative. So the Jacobian matrix is invertible. This implies that for almost all $(\pi, \hat{p})$, the function $g = (g_1, g_2)(\hat{p}, \hat{q}, \pi)$ crosses $(0, 0)$ transversally: there exists a set $\Pi \times P \subset (0, 1) \times (\underline{p}, 1)$ of measure one such that for any $(\pi, \hat{p}) \in \Pi \times P$, the values of q that sustain a symmetric equilibrium satisfy claim (9). ■

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

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

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

■

E Proofs for section 5

E.1 Breakthrough learning

Once a worker generates a breakthrough, his employer keeps him for the rest of time. To track how many workers have "secured their jobs", we let $m(t) \in [0, 1]$ be 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 $\alpha > 1$. Within the next instant, the belief for those matched a-workers who have not generated a breakthrough drops slightly below p_a. Their employers find it optimal to switch to previously unmatched a-workers, the belief for whom is p_a. This is essentially equivalent to all a-workers being employed and allocated $1/\alpha < 1$ of a task at $t = 0$.

In the next instant, those a-workers who have generated a breakthrough stay matched forever (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:

$$\mathrm{d}m(t) = (\alpha - m(t))q(t)\lambda_h \frac{1 - m(t)}{\alpha - m(t)}\mathrm{d}t = q(t)\lambda_h(1 - m(t))\mathrm{d}t \quad \text{and} \quad m(0) = 0. \tag{11}$$

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

$$q(t)(\alpha - m(t)) + m(t) = p_a\alpha \quad \Longrightarrow \quad q(t) = \frac{\alpha p_a - m(t)}{\alpha - m(t)}. \tag{12}$$

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 (11). 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 $\underline{p}$, 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 $\underline{p}$ first, some employers take safe arms.

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

Proof. We first show that as $p_b \uparrow p_a$, $T_b \to 0$. By the definition of T_b and the expression for $q(t)$ in (12), 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 $T_b \to 0$. ■

E.2 Breakdown learning

Under breakdown learning, a matched worker stays matched as long as no breakdown occurs. At time 0, a unit mass of a-workers are matched with employers. When a matched worker generates a breakdown, his employer replaces him with an a-worker who has never been tried before. This process goes on until all the a-workers are tried. From that 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.

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

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

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

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

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

Using the expression we have for $\underline{m}(t)$ 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. ■

F Proofs for section 6

F.1 Stable stage-game matching

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

In this subsection, we characterize the set of stable stage-game matchings for a fixed G. There exists a unique marginal 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 F.1 (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 F.1, 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.

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

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

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

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

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

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

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

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

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

Proof. We prove the first part in two steps.

1. If a less productive worker is matched, then a more productive worker must be matched 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.

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

We now prove the second part. Suppose that a worker's probability of having a high type is $p > p_s$ and he is unmatched. 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 a safe arm. This employer can form a blocking pair with the unmatched worker p. ■

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

Lemma F.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 F.1. 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. ■

F.2 Stable dynamic matching

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

Lemmata F.2 and F.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 F.2;
2. $w^*(p|G) = (p - p^M(G))v$ is the employer-preferred wage function if $w(\cdot)$ is multi-valued as in Lemma F.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 F.1. *Under either breakthrough or breakdown learning, μ^* is dynamically stable.*

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

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

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

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

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

(iii) Suppose that worker i is matched at history h_t according to μ^*. Let $p(t)$ be this worker's probability of 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

$$\max\left\{0, (1 - Q(\tau))\left(p(\tau) - p^M(G(h_\tau))\right)v\right\} = \max\left\{0, p(t)\frac{p(\tau) - p^M(G(h_\tau))}{p(\tau)}v\right\} \tag{13}$$

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.

■

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

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 F.2. *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 E.

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 (13) in the proof of Proposition F.1 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. ■

F.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 (13), 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 E.2, 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 F.3 (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 F.3, 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 F.4 (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

$$\frac{y\left(y^{-p_a} - \frac{x}{y(p_a+x-1)-p_a+1}\right)}{y-1}, \tag{14}$$

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 (14) 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

$$2(p_a + x) < p_a x^2 + 2,$$
$$p_a(x+6)x + 4(x-3)x + 6 < 6p_a,$$

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 F.5 (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 $p_b \uparrow p_a$:

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

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

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. ■

G Proofs for section 7.1

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

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

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

The case of $\lambda_h = \lambda_\ell \geqslant 0$ corresponds to uninformative signals. We ignore this trivial case in the rest of this analysis.

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

Proof. Let $U_i(p_a, p_b)$ be worker i's payoff given the belief pair (p_a, p_b). For any $p_a > p_b$, the 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 $U_a(p_b, p_b) = U_b(p_b, p_b)$. ■

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

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

Proof. Let $U_i(q_a, q_b)$ 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 $r^2 - (1-2p_a)r(\lambda_\ell - \lambda_h) - \lambda_h\lambda_\ell > 0$. ■