Early-Career Discrimination:

Spiraling or Self-Correcting?

Arjada Bardhi[1] Yingni Guo[2] Bruno Strulovici[3]

JHU
November 2023

[1]NYU [23]Northwestern

Early-career discrimination

- Employers know little about the productivity of early-career workers
- They rely on available proxies for workers' productivities
  - observable characteristics (race, gender, ethnicity etc.)
  - Goldin and Rouse (2000), Pager (2003), Bertrand and Mullainathan (2004), Bertrand and Duflo (2016) etc.
  - proxies become less relevant as employer learns from performance

Early-career discrimination

- Employers know little about the productivity of early-career workers
- They rely on available proxies for workers' productivities
  - observable characteristics (race, gender, ethnicity etc.)
  - Goldin and Rouse (2000), Pager (2003), Bertrand and Mullainathan (2004), Bertrand and Duflo (2016) etc.
  - proxies become less relevant as employer learns from performance
- When jobs are scarce, groups discriminated at the start might miss on early opportunities
  - reduced subsequent employment opportunities (Pallais, 2014)
  - reduced access to resources for career advancement (Oyer, 2006)
- Even if groups have very similar productivity distributions

Stylized features

1. Workers from different social groups compete for scarce tasks

Stylized features

1. Workers from different social groups compete for scarce tasks

2. Employers learn through task allocation: a worker's productivity revealed only if the worker performs a task

   belief at date t $\Rightarrow$ allocation at date t

   $\Rightarrow$ belief at date $(t+1)$ $\Rightarrow$ allocation at date $(t+1)$

   $\Rightarrow$ belief at date $(t+2)$ $\Rightarrow$ $\dots$

Stylized features

1. Workers from different social groups compete for scarce tasks

2. Employers learn through task allocation: a worker's productivity revealed only if the worker performs a task

   belief at date t $\Rightarrow$ allocation at date t

   $\Rightarrow$ belief at date $(t+1)$ $\Rightarrow$ allocation at date $(t+1)$

   $\Rightarrow$ belief at date $(t+2)$ $\Rightarrow$ $\ldots$

3. Groups have comparable productivity distributions

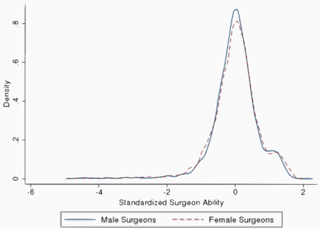
Example: Medical referrals (Sarsons, 2019)

1. Men and women surgeons compete for referrals from physicians
2. Physicians learn about a surgeon's ability only if the surgeon performs a surgery (outcome $\in \{\text{death}, \text{no death}\}$)
3. Men and women surgeons have comparable abilities

   'Women have a lower average ability and a slightly lower variance of ability, but the differences are small.' Sarsons, 2019

FIGURE 2: DISTRIBUTION OF SURGEON ABILITY

(a) Matched Sample



Questions

1. How important is early-career discrimination for workers' lifetime prospects?
2. As groups' productivities converge, do their payoffs converge too?

Questions

1. How important is early-career discrimination for workers' lifetime prospects?
2. As groups' productivities converge, do their payoffs converge too?

Two conjectures on the impact of group belonging:

1. small productivity difference $\rightarrow$ employers learn $\rightarrow$ errors in hiring corrected quickly $\rightarrow$ little impact

Questions

1. How important is early-career discrimination for workers' lifetime prospects?
2. As groups' productivities converge, do their payoffs converge too?

Two conjectures on the impact of group belonging:

1. small productivity difference $\rightarrow$ employers learn $\rightarrow$ errors in hiring corrected quickly $\rightarrow$ little impact
2. small productivity difference $\rightarrow$ unequal early career opportunities $\rightarrow$ different career trajectories $\rightarrow$ significant impact

Questions

1. How important is early-career discrimination for workers' lifetime prospects?

2. As groups' productivities converge, do their payoffs converge too?

Two conjectures on the impact of group belonging:

1. small productivity difference → employers learn → errors in hiring corrected quickly → little impact

2. small productivity difference → unequal early career opportunities → different career trajectories → significant impact

The *nature* of employer learning determines the long-run impact of early-career discrimination

Findings

I. Contrast between learning environments

- Self-correcting environments
  - comparable payoffs to comparable groups
  - such environments track workers' successes
- Spiraling environments
  - large payoff disparities for comparable groups
  - such environments track workers' failures

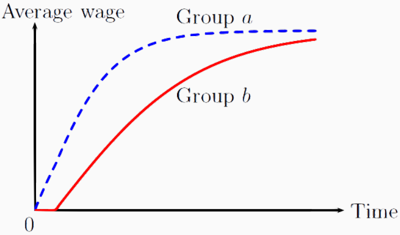
Star jobs vs. guardian jobs (Jacobs, 1981; Baron and Kreps, 1999)

Task scarcity determines how severe spiraling is.

Findings

II. Can flexible wages fix the payoff disparity?

- Contrast persists even with flexible wages
- Both with a zero minimum wage and a strictly negative one
- Comparable groups face very different wage paths
- We analyze the wage gap, employment rate gap, and earnings gap
- Evolution of gaps is consistent with empirical findings



Findings

III. Robustness

1. Worker investment in productivity

   Spiraling gets worse due to polarized investment incentives.

2. Inconclusive learning

   Spiraling persists with sufficiently impatient workers.

3. Misspecified employer beliefs

   Arbitrarily small amounts of misspecification lead to large disparities for identical workers.

Roadmap

Small market

A sharp contrast   Non-vanishing belief difference

Large market

Fixed wages   Flexible wages

Other robustness checks

Investment in productivity   Inconclusive signals
Misspecified prior

Final thoughts

Framework

Players and types

- One employer and two workers: a and b
- Each worker from a distinct social group
- Productivity type of worker i is either high or low: $\theta_i \in \{h, \ell\}$
- Group i's average productivity: $p_i := \Pr(\theta_i = h)$

Comparable social groups

(i) group a has higher productivity: $p_a > p_b$

(ii) groups have almost identical productivity distributions: $p_b \to p_a$

Framework

Task allocation and payoffs

- Continuous time $t \in [0, \infty)$, long-lived players, discount $r > 0$
- At each t, employer allocates a divisible task

$$\{ \text{ worker } a, \text{ worker } b, \text{ safe arm } \}$$

- Employer's flow payoff:
    - $v > 0$ if task goes to a high-type worker
    - 0 if task goes to a low-type worker
    - $s \in (0, v)$ if safe arm
- Worker's flow payoff:
    - fixed wage $w = 1$ if allocated the task
    - 0 otherwise

Employer's problem is a standard three-armed bandit problem.

Framework

Learning environment

worker i is allocated the task over $[t, t + \mathrm{d}t)$

$$\Downarrow$$

employer learns about θ_i over $[t, t + \mathrm{d}t)$

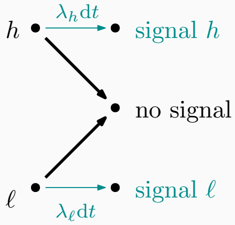
Learning environment

Learning environment described by a pair of Poisson arrival rates

$$(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$$

If worker of type θ allocated the task over $[t, t + \mathrm{d}t)$

- conclusive public news arrives with probability $\lambda_\theta \, \mathrm{d}t$
- no news arrives with probability $1 - \lambda_\theta \, \mathrm{d}t$



Framework

Learning environment

Learning environment described by a pair of Poisson arrival rates

$$(\lambda_h, \lambda_\ell) \in \mathbb{R}_+^2$$

If worker of type θ allocated the task over $[t, t + \mathrm{d}t)$

- conclusive public news arrives with probability $\lambda_\theta \, \mathrm{d}t$
- no news arrives with probability $1 - \lambda_\theta \, \mathrm{d}t$

Two classes of learning environments:

1. breakthrough-salient: $\lambda_h > \lambda_\ell \geqslant 0$
2. breakdown-salient: $\lambda_\ell > \lambda_h \geqslant 0$

If $\lambda_\ell = \lambda_h$, no news is entirely uninformative

Learning environment

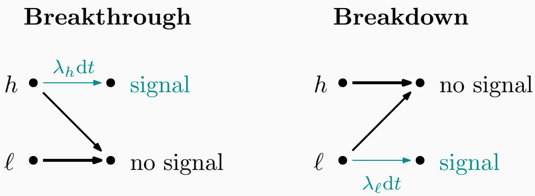
Learning environment described by a pair of Poisson arrival rates

$$(\lambda_h, \lambda_\ell) \in \mathbb{R}^2_+$$

If worker of type θ allocated the task over $[t, t + \mathrm{d}t)$

- conclusive public news arrives with probability $\lambda_\theta \, \mathrm{d}t$
- no news arrives with probability $1 - \lambda_\theta \, \mathrm{d}t$

Pure learning environments:



Interpreting learning environments

- Intrinsic feature of the job considered
- Tracking under-performance (breakdowns) vs. over-performance (breakthroughs)
- Jacobs (1981), Baron and Kreps (1999):

"star jobs" vs. "guardian jobs"

Interpreting learning environments

- Intrinsic feature of the job considered
- Tracking under-performance (breakdowns) vs. over-performance (breakthroughs)
- Jacobs (1981), Baron and Kreps (1999):

"star jobs" vs. "guardian jobs"

'The first-rate salesman can often add a significant increment to the performance of his organization while his inferior will not impose unacceptable costs.' Jacobs, 1981

Interpreting learning environments

- Intrinsic feature of the job considered
- Tracking under-performance (breakdowns) vs. over-performance (breakthroughs)
- Jacobs (1981), Baron and Kreps (1999):

"star jobs" vs. "guardian jobs"

'The airline pilot who misses a landing or the operative who inadvertently blocks a long assembly line will produce rather destructive effects, but an outstanding performance in either position will be of little consequence for the organization.' Jacobs, 1981

Baron and Kreps (1999)

Breakthrough vs. breakdown learning

1. Do workers' lifetime payoffs converge as $p_b \uparrow p_a$?
2. Which learning environments, if any, grant a disproportionate first-hire advantage?

Related work

1. **Statistical discrimination**:
   - Phelps (1972), Aigner and Cain (1977), Cornell and Welch (1996)
   - Arrow (1973), Foster and Vohra (1992), Coate and Loury (1993)
2. **Employer learning**: Farber and Gibbons (1996), Altonji and Pierret (2001), Altonji (2005), Lange (2007), Antonovics and Golan (2012), Mansour (2012), Pallais (2014), Bose and Lang (2017)
3. **Poisson bandit approach**: Bergemann and Valimaki (1996), Felli and Harris (1996), Keller, Rady, and Cripps (2005), Strulovici (2010), Keller and Rady (2010, 2015), Che and Hörner (2018), Board and Meyer-ter-Vehn (2013), Halac and Kremer (2020), Lizzeri, Shmaya, Yariv (2021)
4. **Experimentation + discrimination**: Fershtman and Pavan (2020), Komiyama and Noda (2020), Li, Raymond and Bergman (2020), Che, Kim and Zhong (2019)

Breakthrough learning

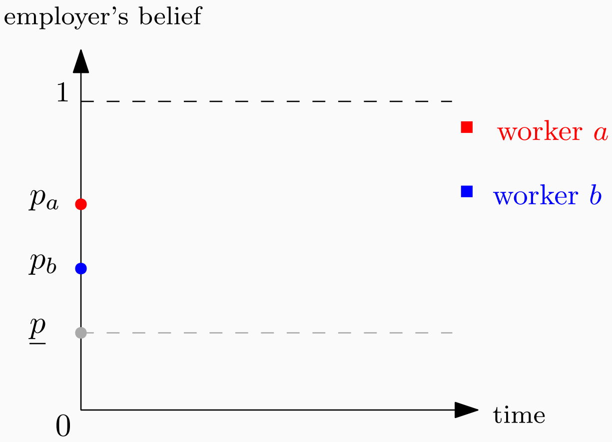
Optimal allocation

At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.

Breakthrough learning

Optimal allocation

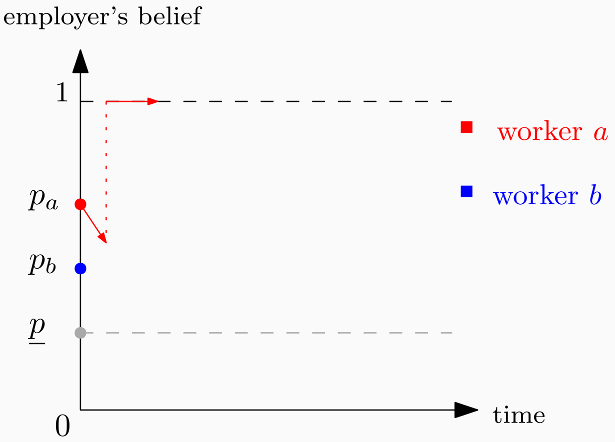
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakthrough learning

Optimal allocation

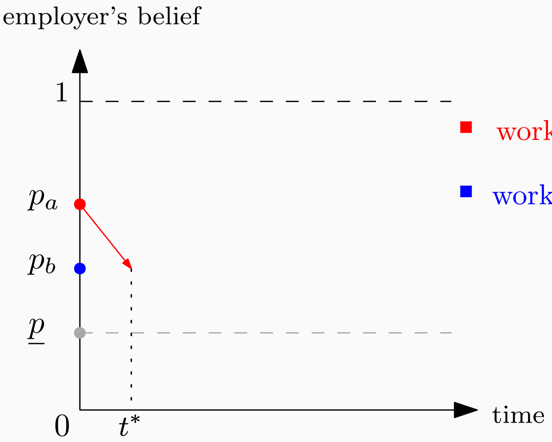
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakthrough learning

Optimal allocation

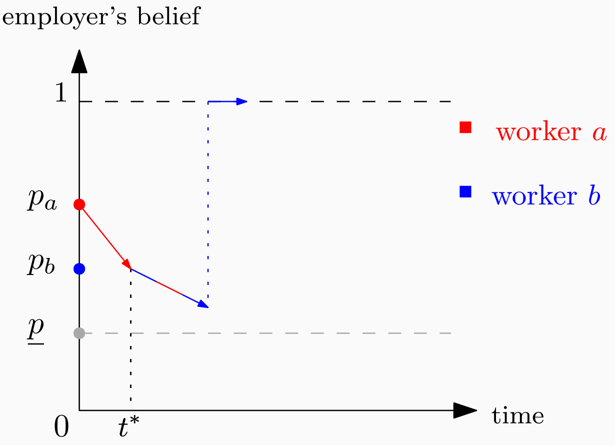
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakthrough learning

Optimal allocation

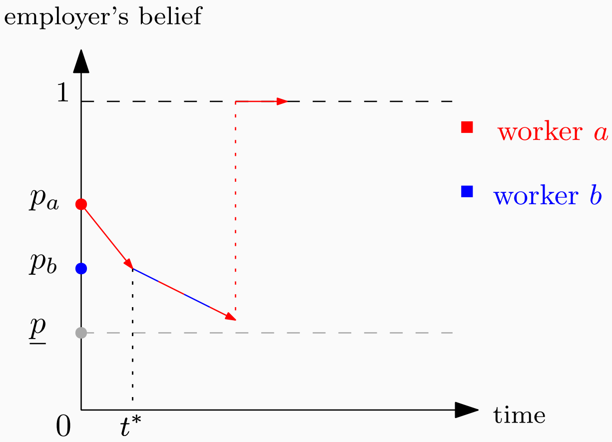
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakthrough learning

Optimal allocation

At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Self-correction in breakthrough learning

Proposition 1a

Let $\lambda_h > \lambda_\ell \geqslant 0$. As $p_b \uparrow p_a$, the expected payoff of worker b converges to that of worker a.

- beliefs drift down in the absence of news
- task assigned exclusively to worker a over $[0, t^*]$

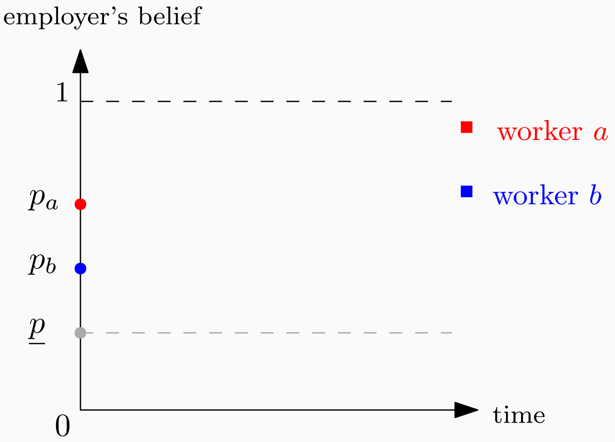
$$t^* = \frac{1}{\lambda_h - \lambda_\ell} \log\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right)$$

  unless worker a generates a breakdown
- workers treated symmetrically after t^*
- as $p_b \uparrow p_a$, grace period $t^* \to 0$
- the advantage of worker a vanishes

Breakdown learning

Optimal allocation

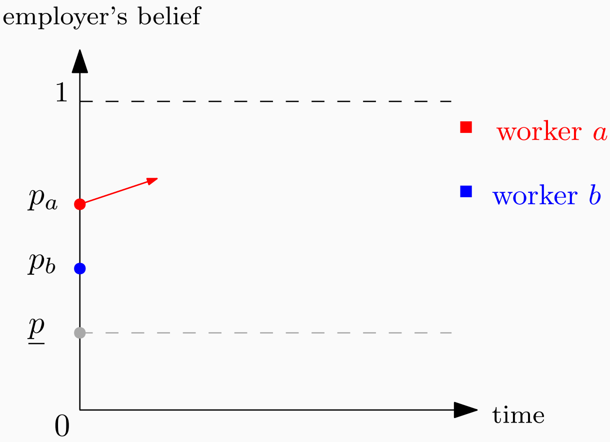
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakdown learning

Optimal allocation

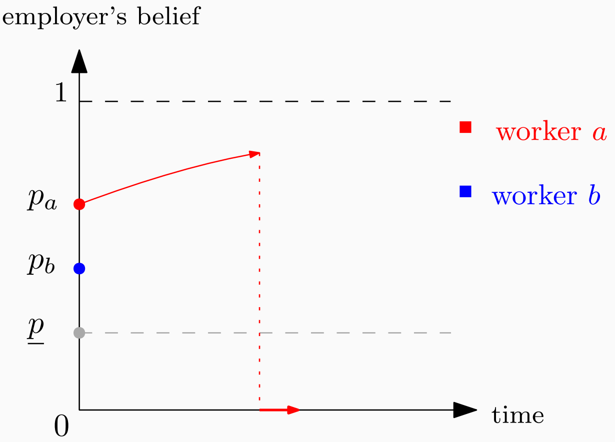
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakdown learning

Optimal allocation

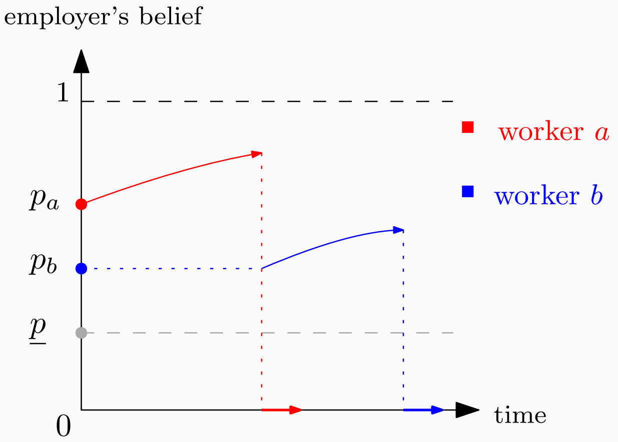
At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Breakdown learning

Optimal allocation

At each t, employer allocates the task to either the worker that is more likely to be h or the safe arm ($\underline{p}$), where $\underline{p}$ depends on $\max\{\lambda_\ell, \lambda_h\}$.



Spiraling in breakdown learning

Proposition 1b

Let $\lambda_\ell \geqslant \lambda_h \geqslant 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a converges to

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

Spiraling in breakdown learning

Proposition 1b

Let $\lambda_\ell \geqslant \lambda_h \geqslant 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a converges to

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

- belief drifts up in the absence of news
- task assigned to worker a until he realizes a breakdown
- worker a's payoff

$$\underbrace{p_a}_{\text{no breakdown ever}} + (1 - p_a) \cdot \underbrace{\frac{r}{\lambda_\ell + r}}_{\text{expected time until a breakdown}}$$

- worker b's payoff

$$\underbrace{(1 - p_a)\frac{\lambda_\ell}{\lambda_\ell + r}}_{b \text{ gets a chance}} \left(p_b + (1 - p_b)\frac{r}{\lambda_\ell + r} \right)$$

Spiraling in breakdown learning

Proposition 1b

Let $\lambda_\ell \geqslant \lambda_h \geqslant 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a converges to

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

- order effect: if $p_a, p_b \approx 1$, worker b never gets a chance
- delay effect: if $p_a, p_b \approx 0$, worker b is very likely to get a chance but with substantial delay if λ_ℓ small
- gap persists even for very fast learning: $(1 - p_a)$ as $\lambda_\ell \to +\infty$
- but gap becomes smaller with higher λ_ℓ: first-order effect is to reduce delay for worker b

Spiraling in breakdown learning

Proposition 1b

Let $\lambda_\ell \geqslant \lambda_h \geqslant 0$. As $p_b \uparrow p_a$, the ratio of the expected payoff of worker b to that of worker a converges to

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

- in fact spiraling arises even in the symmetric environment $\lambda_h = \lambda_\ell$
- task allocation dynamics and workers' payoffs are the same as in any breakdown-salient environment with $(\lambda_\ell, \cdot)$

Contrast between breakthrough and breakdown learning

As $p_b \uparrow p_a$, worker a's advantage from early-career discrimination:

Contrast between breakthrough and breakdown learning

As $p_b \uparrow p_a$, worker a's advantage from early-career discrimination:

- vanishes in breakthrough-salient environment
  - comparable workers $\Rightarrow$ comparable lifetime payoffs
  - downward belief drift leads to quick equalization of comparable workers

Contrast between breakthrough and breakdown learning

As $p_b \uparrow p_a$, worker a's advantage from early-career discrimination:

- vanishes in breakthrough-salient environment
  - comparable workers $\Rightarrow$ comparable lifetime payoffs
  - downward belief drift leads to quick equalization of comparable workers
- persists in breakdown-salient environments
  - comparable workers $\not\Rightarrow$ comparable lifetime payoffs
  - upward belief drift goes against equalization of comparable workers

Small belief difference

A small $(p_a - p_b)$ rather than $p_b \uparrow p_a$ reveals a more nuanced picture

Fix $\lambda_h + \lambda_\ell =$ constant

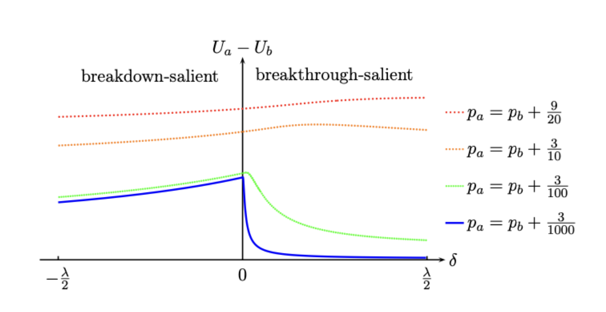


Figure 1: $\lambda = 2, s = 1/10, v = r = 1, p_b = 1/3$

The payoff gap is continuous in the learning environment.

Roadmap

Small market

A sharp contrast   Non-vanishing belief difference

Large market

Fixed wages   Flexible wages

Other robustness checks

Investment in productivity   Inconclusive signals
Misspecified prior

Final thoughts

Framework

- Unit mass of tasks, α mass of *a*-workers, β mass of *b*-workers
- Task scarcity: more workers than tasks
  - $\alpha + \beta > 1$, but for exposition suppose $\alpha > 1$

Framework

- Unit mass of tasks, α mass of *a*-workers, β mass of *b*-workers
- Task scarcity: more workers than tasks
  - $\alpha + \beta > 1$, but for exposition suppose $\alpha > 1$
- A stage-game matching specifies
  (i) how workers are matched to employers
  (ii) a (non-negative) wage for each matched pair
- Public learning: employers share the same belief about each worker

Framework

- Unit mass of tasks, α mass of *a*-workers, β mass of *b*-workers
- Task scarcity: more workers than tasks
  - $\alpha + \beta > 1$, but for exposition suppose $\alpha > 1$
- A stage-game matching specifies
  - (i) how workers are matched to employers
  - (ii) a (non-negative) wage for each matched pair
- Public learning: employers share the same belief about each worker
- Shapley and Shubik (1971): A stage-game matching is *stable* if
  - (i) no matched employer *j* strictly prefers to take a safe arm instead,
  - (ii) there exists no blocking pair.
- We characterize the essentially unique stable stage-game matching
- Then we show that repeating the stable stage-game matching after each history μ^* is dynamically stable

Framework

Definition (Ali and Liu, 2020)

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, such that

(i) no matched employer j under $\mu|_{h_t}$ who strictly prefers to take the 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 $\mu|_{h_{t+\mathrm{d}t}}$;

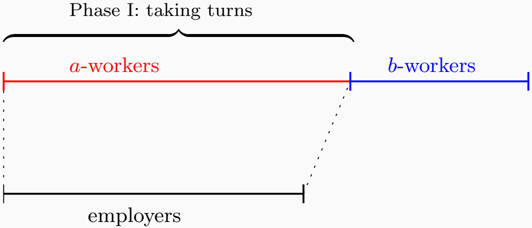
(iii) no matched worker i under $\mu|_{h_t}$ who strictly prefers to be unmatched over $[t, t + \mathrm{d}t)$ and then revert to $\mu|_{h_{t+\mathrm{d}t}}$.

Proposition

Under both pure learning environments, μ^* is dynamically stable.

Large market with fixed wages

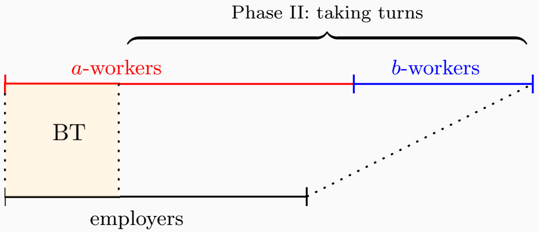
Diverse hiring in the pure breakthrough environment



Phase I lasts until belief of a-workers drops to p_b.

Large market with fixed wages

Diverse hiring in the pure breakthrough environment



Phase II lasts until either all employers obtain a breakthrough or belief hits $\underline{p}$.

Large market with fixed wages

Diverse hiring in the pure breakthrough environment

Phase I : tasks split between *a*-workers only

Phase II : remaining tasks split between *a*-workers and all *b*-workers

Large market with fixed wages

Diverse hiring in the pure breakthrough environment

Phase I : tasks split between *a*-workers only

Phase II : remaining tasks split between *a*-workers and all *b*-workers

Self-correction under breakthroughs

Delay for group *b* vanishes as $p_b \uparrow p_a$. Hence, *a*-workers and *b*-workers have the same expected payoff.

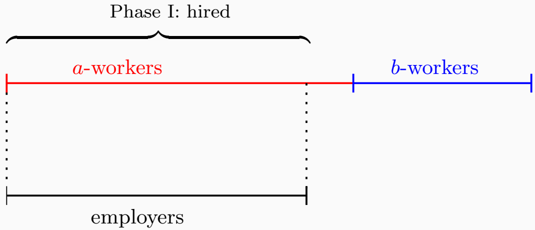
Large market with fixed wages

Diverse hiring in the pure breakthrough environment

> *'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

Large market with fixed wages

Narrow hiring in the pure breakdown environment



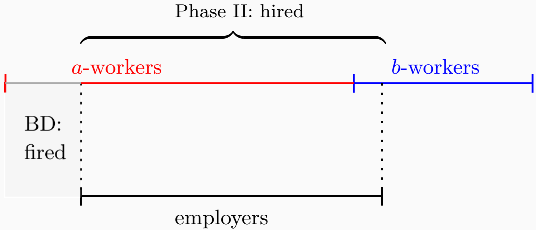
Phase I lasts until at least $(\alpha - 1)$ a-workers generate breakdowns.

Large market with fixed wages

Narrow hiring in the pure breakdown environment

Large market with fixed wages

Narrow hiring in the pure breakdown environment



Phase II lasts until all employers obtain a breakdown (potentially never!).

Large market with fixed wages

Narrow hiring in the pure breakdown environment

Phase I : *a*-workers hired only

Phase II : *b*-workers hired after sufficiently many *a*-workers failed

Large market with fixed wages

Narrow hiring in the pure breakdown environment

Phase I : a-workers hired only

Phase II : b-workers hired after sufficiently many a-workers failed

Spiraling under breakdowns

Delay for group b does not vanish as $p_b \uparrow p_a$. So b-workers obtain a strictly lower expected payoff than a-workers.

How does task scarcity affect inequality under spiraling?

Proposition (Inequality increases in task scarcity)

As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker decreases in both α and β.

How does task scarcity affect inequality under spiraling?

Proposition (Inequality increases in task scarcity)

As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker decreases in both α and β.

- $\beta \uparrow$: intensifies competition among *b*-workers but no effect on *a*-workers
- $\alpha \uparrow$: intensifies competition among *a*-workers and uniformly delays hiring of *b*-workers
    - *b*-workers are hurt more than *a*-workers

How does task scarcity affect inequality under spiraling?

Proposition (Inequality increases in task scarcity)

As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker decreases in both α and β.

- $\beta \uparrow$: intensifies competition among *b*-workers but no effect on *a*-workers
- $\alpha \uparrow$: intensifies competition among *a*-workers and uniformly delays hiring of *b*-workers
  - *b*-workers are hurt more than *a*-workers

While all groups suffer during economic downturns, some suffer disproportionately more.

How does task scarcity affect inequality under spiraling?

Proposition (Inequality increases in task scarcity)

As $p_b \uparrow p_a$, the limiting ratio of the expected payoff of a b-worker to that of an a-worker decreases in both α and β.

- $\beta \uparrow$: intensifies competition among *b*-workers but no effect on *a*-workers
- $\alpha \uparrow$: intensifies competition among *a*-workers and uniformly delays hiring of *b*-workers
  - *b*-workers are hurt more than *a*-workers

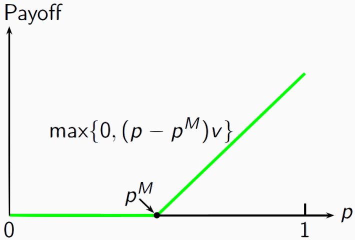
While all groups suffer during economic downturns, some suffer disproportionately more.

But can flexible wages fix spiraling?

- No, as long as both groups have the same ability to accept negative wages and the minimum wage is not too negative.

Stable stage-game matching with flexible wage

- Fix a time t and for each $p \in [0,1]$, let $G(p)$ denote the fraction of workers i s.t. $p_i \leq p$
- There is a G-dependent marginal productivity cutoff p^M



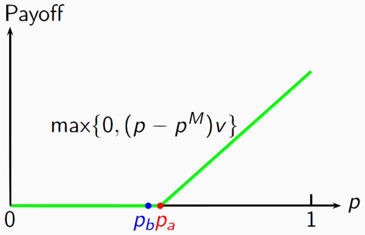
- Worker is matched iff his expected productivity p exceeds p^M
- Earnings: $\left(p - p^M\right) v$ for matched workers and 0 for unmatched ones
- All employers make the same expected profit
- $\mu^* =$ fixed-wage dynamic allocation + this wage

Can flexible wages fix spiraling?

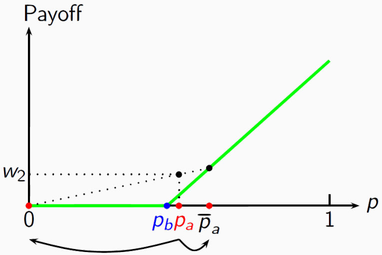
- The worker's static payoff is convex
- More learning about a worker's type $\Rightarrow$ higher expected payoff
- Delay for group b does not vanish as $p_b \uparrow p_a$
- More is learned about a-workers than b-workers

Can flexible wages fix spiraling?

- The worker's static payoff is convex
- More learning about a worker's type ⇒ higher expected payoff
- Delay for group b does not vanish as $p_b \uparrow p_a$
- More is learned about a-workers than b-workers
- Consider a two-period intuition with $\alpha = \beta = 1$



First period



Second period

Persistent wage and earnings gaps

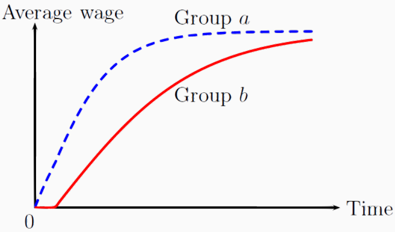
Proposition

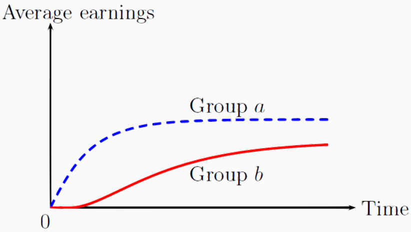
In μ^*, as $p_b \to p_a$ the average lifetime earnings of a-workers converge to those of b-workers under breakthroughs but not under breakdowns.

Persistent wage and earnings gaps

Proposition

In μ^*, as $p_b \to p_a$ the average lifetime earnings of a-workers converge to those of b-workers under breakthroughs but not under breakdowns.





- Earnings gap = wage gap + employment rate gap
- If task scarcity too severe (more high-type workers than tasks), non-zero earnings gap even as $t \to \infty$

Strictly negative minimum wage

If *b*-workers are willing to accept a negative wage, *a*-workers will outbid down to minimum wage

Revise the dynamic matching μ^* by lowering marginal worker's wage to $\underline{w}$

$$w_t(p_i) = (p_i - p^M(G(h_t)))v + \underline{w}$$

Dynamic allocation of tasks remains the same as before

Strictly negative minimum wage

If *b*-workers are willing to accept a negative wage, *a*-workers will outbid down to minimum wage

Revise the dynamic matching μ^* by lowering marginal worker's wage to $\underline{w}$

$$w_t(p_i) = (p_i - p^M(G(h_t)))v + \underline{w}$$

Dynamic allocation of tasks remains the same as before

Proposition

Let $\alpha p_a < 1 < \alpha p_a + \beta p_b$. As $p_b \to p_a$, $\mu^*(\underline{w})$ is dynamically stable in the breakdown environment for any

$$\underline{w} > -\frac{\lambda(1 - p_b)p_b v}{\lambda p_b + r},$$

which is equivalent a b-worker's continuation payoff at time 0 being strictly positive.

Stylized facts

- Statement on Gender Salary Equity

  'The 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.'

  Association of Women Surgeons (2017)

Stylized facts

- Statement on Gender Salary Equity

  > *'The 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.'*
  >
  > Association of Women Surgeons (2017)

- Arcidiacono (2003), Arcidiacono, Bayer and Hizmo (2010): racial earnings/wage gaps are small at early career stages but widen with labor market experience

Stylized facts

- Statement on Gender Salary Equity

  'The 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.'

  Association of Women Surgeons (2017)

- Arcidiacono (2003), Arcidiacono, Bayer and Hizmo (2010): racial earnings/wage gaps are small at early career stages but widen with labor market experience

- Low-skill jobs ≈ breakdown / high-skill jobs ≈ breakthrough

  'Thus, while one challenge is to explain earning differentials between black and white men, there is an even greater challenge, which is to explain the simultaneous existence of wage differentials among relatively low-skill male workers and their possible absence among high-skill male workers.' Lang and Lehmann (2012)

Roadmap

Small market

A sharp contrast   Non-vanishing belief difference

Large market

Fixed wages   Flexible wages

Other robustness checks

Investment in productivity   Inconclusive signals
Misspecified prior

Final thoughts

Opportunity to invest

- Back to our two-worker one-employer baseline model
- Before $t = 0$, each ℓ-type worker draws his investment cost from distribution F on $[0, 1]$, and decides whether to invest
- If a worker invests, his type improves to h w.p. $\pi \in [0, 1]$
- The pre-investment and post-investment types are private information to the worker
- F is the same for both groups

Investment equilibrium

- Let (q_a, q_b) denote the employer's post-investment belief pair
- The employer chooses an optimal allocation strategy given (q_a, q_b)
- Worker i's benefit from investment is

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

  where $U_i(\theta_i; q_a, q_b)$ is the expected payoff of type θ_i of worker i given (q_a, q_b)
- Worker i invests if and only if his realized cost is below $B_i(q_a, q_b)$
- An equilibrium is a pair of beliefs (q_a, q_b) such that for both $i = a, b$

$$q_i = p_i + (1 - p_i)\pi F\left(B_i(q_a, q_b)\right)$$

Multiplicity in investment equilibria

- (Post-investment) favored worker has stronger incentives to invest than the discriminated one
- This self-fulfilling force leads to multiple investment equilibria
- There exist equilibria in which b overtakes a and becomes favored

Equilibrium sets

We compare the equilibrium sets across the two pure learning environments.

Findings

Result 1: equilibrium lifetime payoffs for workers

- Investment does not disturb the self-correcting property of breakthroughs
- Investment exacerbates spiraling under breakdowns: it makes the workers' payoffs more unequal than without investment

Result 2: workers' investment behavior

- With sufficiently fast learning, breakdown learning leads to more polarized investment across workers than breakthrough learning
- discriminated BD $<$ discriminated BT $<$ favored BT $<$ favored BD
- when $\pi \approx 1$ also, in the breakdown environment favored worker invests and is a high type almost surely $\Rightarrow$ employer prefers BD

Self-correcting property

Investment does not disturb the self-correcting property of breakthroughs

Proposition

As $p_b \uparrow p_a$, there exists an equilibrium in which the two workers' expected payoffs as well as their post-investment beliefs converge.

- If $p_a = p_b$, there exists a symmetric equilibrium in which $q_a = q_b$
- Benefit $B_i(q_a, q_b)$ is continuously differentiable in (q_a, q_b)
- By the implicit function theorem, as $p_b \uparrow p_a$, there exists an equilibrium in which $q_a \uparrow q_b$

Exacerbated spiraling under breakdowns

Investment exacerbates spiraling under breakdowns

Proposition

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 $(1 - p_a)\lambda_\ell/(\lambda_\ell + r)$.

- Payoff ratio is determined by how likely it is that the favored worker has a high type
- With investment, favored worker is more likely to be a high type
- So inequality is higher too

Final thoughts

Inconclusive signals

Suppose a single signal can be generated by both types

If a worker of type $\theta \in \{h, \ell\}$ is allocated the task over $[t, t + \mathrm{d}t)$

- public signal arrives with probability $\lambda_\theta \, \mathrm{d}t$
- no signal arrives with probability $1 - \lambda_\theta \, \mathrm{d}t$

Two classes of learning environments

1. breakthrough: $\lambda_h > \lambda_\ell \geqslant 0$
2. breakdown: $\lambda_\ell > \lambda_h \geqslant 0$

Inconclusive breakthroughs

Proposition

For any $\lambda_h > \lambda_\ell$, the two workers' payoffs converge as $p_b \to p_a$.

- If no signal until time $t^*(p_a, p_b)$, principal splits the task between a and b afterwards
- $t^*(p_a, p_b)$ converges to zero as $p_b \uparrow p_a$
- Hence, payoffs converge

Inconclusive breakdowns

Proposition

For any $\lambda_h < \lambda_\ell$ the ratio of the two workers' payoffs is bounded away from 1 for sufficiently impatient players:

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

- Intuition: time until first breakdown sufficiently long
  - first-hire advantage substantial
  - due to impatience, time to the first breakdown dominates

- Alternatively, arrival rates λ_ℓ and λ_h are low (infrequent breakdowns)

Misspecified prior belief

- Objective expected productivities are $p_a = p_b = p_{true}$
- Employer believes that $\hat{p}_b = p_{true}$ but $\hat{p}_a = p_{true} + \epsilon$
- Worker a is hired first in any learning environment
- Self-correcting property of breakthroughs still holds
  - As misspecification vanishes $\epsilon \downarrow 0$, so does advantage t^*
- However spiraling persists even though groups are identical
  - Vanishing misspecification breaks the tie at $t = 0$
  - As $\epsilon \downarrow 0$, the ratio of payoffs goes to $(1 - p_{true})\frac{\lambda_\ell}{\lambda_\ell + r}$
- Lang and Lehmann (2012) argue that mild prejudice among employers is widespread
  - even very mild prejudice harmful in breakdown-salient environments

Roadmap

Small market

A sharp contrast   Non-vanishing belief difference

Large market

Fixed wages   Flexible wages

Other robustness checks

Investment in productivity   Inconclusive signals
Misspecified prior

Final thoughts

Matthew effect

Our results propose a novel mechanism behind the Matthew effect

Merton (1968): initial (dis)advantage begets further (dis)advantage

- Merton observed this in the context of scientific recognition and cumulative privilege
- Scientists of established reputation receive disproportionate credit for either joint or simultaneous discoveries, which further advances their reputation
- Scarcity plays a key role:
  - "the phenomenon of the 41st chair is an artifact of having a fixed number of places available at the summit of recognition"
  - French Academy of Science's practice of restricting the number of members to only 40 scientists at any time

Final thoughts

> *'How economically relevant statistical discrimination is depends on how fast employers learn about workers' productive types'*
>
> Lange (2007)

- The nature of learning – not just the speed – is key for early-career discrimination.
- Early-career discrimination among comparable workers can have a significant lifetime impact
- More empirical work needed on the persistence and magnitude of discrimination in star vs. guardian jobs

Thank you

Performance distribution

- How to empirically test whether a job is a breakthrough-salient job or a breakdown-salient one
- A right-skewed density of performance suggests a BT while a left-skewed density suggests a BD job
- O'Boyle and Aguinis (2012) and Aguinis and Bradley (2015): in occupations centered around star performance (researchers, entertainers, athletes) the distribution is right-skewed

Back

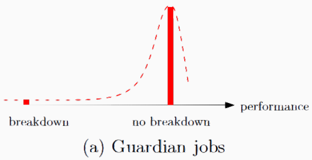


Figure 2: Distribution of performance (adapted from Baron and Kreps (1999))