Complementary Bias: A Model of Two-Sided Statistical Discrimination

Ashley C. Craig and Roland G. Fryer, Jr.

NBER Working Papers 23811, National Bureau of Economic Research, Inc

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Winter 2021

1. Introduction

2. A Brief Review of the Literature

A. Models of Discrimination

B. Strategic Complementarities

3. The Basic Model

A. Building Blocks

- Imagine a large number of employers and a larger population of workers.
- Each employer is randomly matched with many workers from this population.
- Workers belong to one of two identifiable groups, $j \in \{A, B\}$
- Denote by λ_A the fraction of *A*s in the population and $\lambda_B = 1 - \lambda_A$ the fraction of *B*s.
- One can imagine groups being race, gender, or any other protected class.

- Nature moves first and assigns a type to each worker and a type to each employer.
- The worker's type, denoted by $c \in (0, \bar{c}), \bar{c} < \infty$, depicts her cost of investment in human capital.
- Let the fraction of workers with costs no greater than c be represented by $G^W(c)$ – a smooth and continuous cumulative distribution function – with $g^W(c)$ the associated density.
- Similarly, employers have the opportunity to invest at a cost $k_j \in \left(o, \bar{k}\right), \bar{k} < \infty$, to make their workplaces desirable and productive places to work for workers of type j.

- The fraction of employers with investment cost no greater than k_j is $G^E(k_j)$, with $g^E(k_j)$ the associated density.

- Superscripts "*W*" and "*E*" refer to workers and employers, respectively.

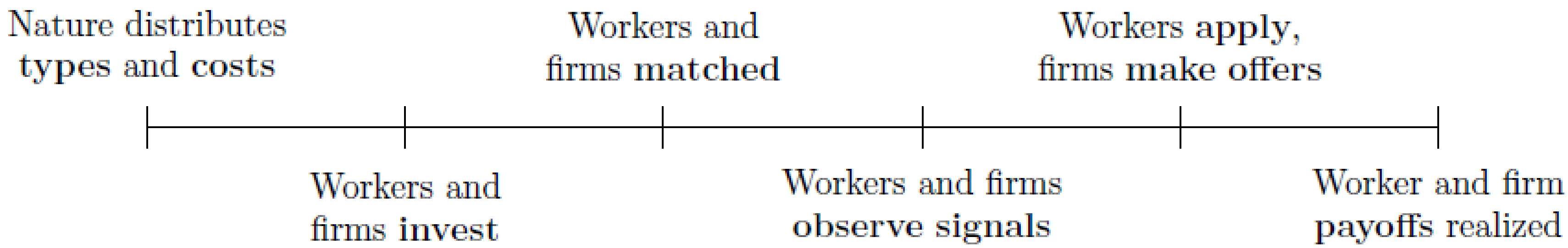


Figure 1: Sequence of Actions

- Associated smooth and continuous distribution function $F_i^W(\theta)$
- Density function $f_i^W(\theta)$ where $i \in \{q, u\}$
- $\phi(\theta) \equiv \frac{f_u^W(\theta)}{f_q^W(\theta)}$ is non-increasing in θ (i.e., $f_i^W(\theta)$ satisfies the monotone likelihood ration property)
- Noisy but informative signal $\psi \in [0,1]$ to workers
- Distribution function $F_i^E(\psi)$
- Density function $f_i^E(\psi)$ where $i \in \{q, u\}$

B. Payoffs

- If the worker is hired and works for an employer who has made a group j investment, she receives a payoff of $\omega_q - c$ if she chose to invest

- ω_q if not.

- If the worker is hired and works for an employer who has not made a group j investment, she receives $-\omega_u - c$ if she invested and $-\omega_u$ if she did not

- If she does not work for any employer, she receives $-c$ if she invested or zero otherwise

C. Strategies

- The worker's strategy consists of a pair of functions – an investment decision and an application decision

- Write as $I^W\colon\{A,B\} \times [0,\bar{c}] \rightarrow [0,1]$ and $A^W\colon\{A,B\} \times [0,1\,] \times [0,1] \times [0,\bar{c}] \rightarrow [0,1]$

- The employer's strategy also consists of a pair of functions – an investment decision and an assignment decision $-I^E\colon\{A,B\} \times \left[0,\bar{k}\right] \rightarrow [0,1], A^E\colon\{A,B\} \times [0,1] \times [0,1] \times \left[0,\bar{k}\right] \rightarrow [0,1]$

D. Expected Payoffs

- Investing in human capital increases the likelihood that a worker is accepted by an employer.
- If a worker of type j invests, she gets expected gross payoff $(1 - F_q^W(s_j))\ \overline{\omega}\ (\delta_j)$
- Conversely, if she does not invest, she gets $(1 - F_u^W(s_j))\ \overline{\omega}\ (\delta_j)$
- Thus, the net return on investment for workers can be written as:

$$\beta_W\left(s_j, \delta_j\right) \equiv \left[F_u^W\left(s_j\right) - F_q^W\left(s_j\right)\right] \overline{\omega}\left(\delta_j\right). \tag{1}$$

The net return on investment for workers can be written as:

$$\beta_E(t_j, \pi_j | \lambda_j) \equiv \lambda_j \left[F_u^E(t_j) - F_q^E(t_j)\right] \overline{\chi}(\pi_j) \,. \qquad (2)$$

E. Bayesian Nash Equilibrium

Definition. *An equilibrium of the game is a pair of beliefs* $\{\pi, \delta\}$ *satisfying:*

$$\pi_j = G^W\left(\beta_W\left(s_j, \delta_j\right)\right) \tag{3}$$

$$\delta_j = G^E\left(\beta_E\left(t_j, \pi_j | \lambda_j\right)\right) \tag{4}$$

Proposition 1. *Let* $\pi^*(\delta)$ *and* $\delta^*(\pi)$ *be the sets of solutions to equations (3) and (4) respectively. Assume that* $\phi(\theta)$ *and* $\tau(\psi)$ *are continuous, strictly decreasing and strictly positive on* $[0,1]$, *and that* $G^W(c)$ *and* $G^E(k)$ *are continuous with full support on* $[0,\bar{c}]$ *and* $[0,\bar{k}]$ *with* $G^W(0) = G^E(0) = 0$. *Further assume that for some* $\underline{\delta}$, *there exists an* s *for which* $G^W(\beta_W(s,\delta)) > \phi(s)/[\chi_q/\chi_u + \phi(s)]$. *Similarly assume that for some* $\underline{\pi}$, *there exists a* t *for which* $G^E(\beta_E(t,\pi|\lambda)) > \tau(t)/[\omega_q/\omega_u + \tau(t)]$. *Then non-zero elements of* $\pi^*(\delta)$ *and* $\delta^*(\pi)$ *exist for any* $\delta \geq \underline{\delta}$ *and* $\pi \geq \underline{\pi}$ *respectively. If there is a set of beliefs* $\{\pi,\delta\}$ *such that* $\delta \in \delta^*(\pi)$ *and* $\pi < \max\{\pi^*(\delta)\}$ *then there exist multiple solutions to the two-sided model.*

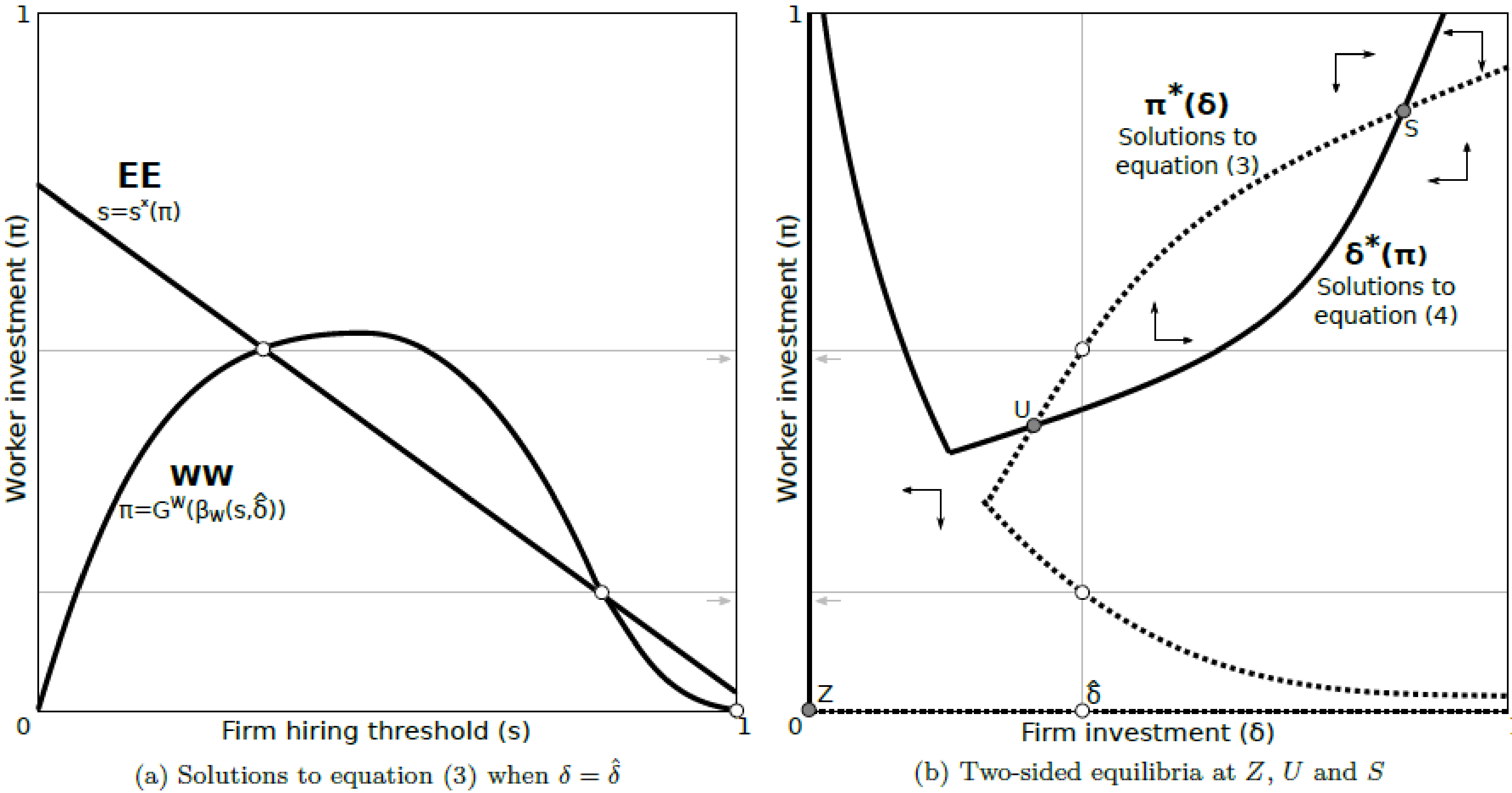


Figure 2: Equilibria in the two-sided model

F. Dynamics

To analyze stability more generally, we approximate this two-dimensional system of non-linear equations as a first-order linearized system of difference equations.

$$d_{t+1} = \begin{pmatrix} \pi_{t+1} \\ \delta_{t+1} \end{pmatrix} = \Upsilon \begin{pmatrix} \pi_t \\ \delta_t \end{pmatrix} \tag{5}$$

For ease of exposition, define the following derivatives.

$$WW_1' = G^{W\prime} \cdot \left[f_u^W\left(s^*\left(\pi\right)\right) - f_q^W\left(s^*\left(\pi\right)\right)\right] \qquad WW_2' = G^{E\prime} \cdot \left[f_u^E\left(t^*\left(\delta\right)\right) - f_q^E\left(t^*\left(\delta\right)\right)\right]$$

$$EE_1' = 1/s^{*\prime}\left(\pi\right) \qquad EE_2' = 1/t^{*\prime}\left(\delta\right)$$

$$RR_1' = \overline{\omega}'\left(\delta\right) \cdot \left[F_u^W\left(s^*\left(\pi\right)\right) - F_q^W\left(s^*\left(\pi\right)\right)\right] \cdot G^{W\prime} \qquad RR_2' = \overline{\chi}'\left(\pi\right) \cdot \lambda \cdot \left[F_u^E\left(t^*\left(\delta\right)\right) - F_q^E\left(t^*\left(\delta\right)\right)\right] \cdot G^{E\prime}$$

These definitions allow us to write the Jacobian of the system compactly.

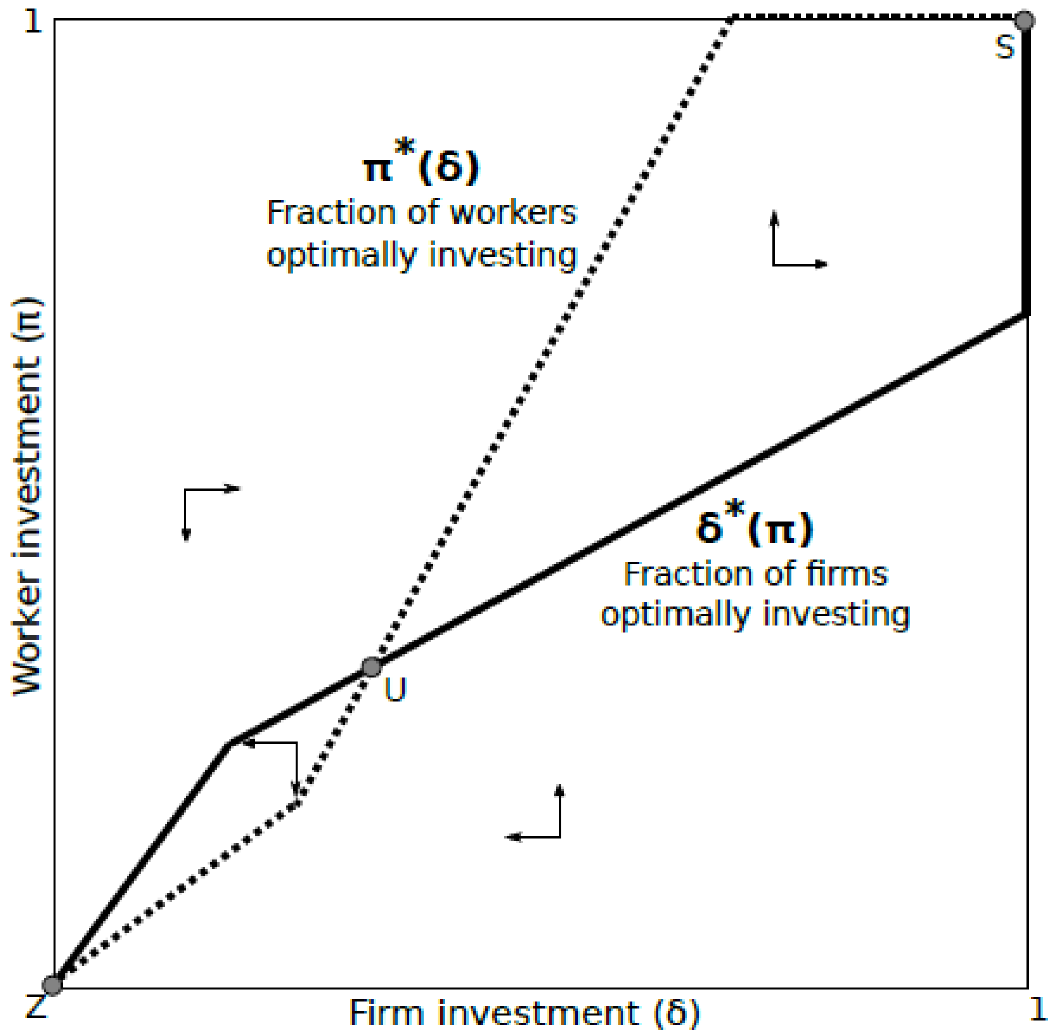
$$\begin{bmatrix} WW_1' \frac{1}{EE_1'} & RR_1' \\ RR_2' & WW_2' \frac{1}{EE_2'} \end{bmatrix}$$

The following condition is necessary and sufficient for this (Neusser 2016).

$$\left|WW_1'\frac{1}{EE_1'}+WW_2'\frac{1}{EE_2'}\right|<1+\left(WW_1'\frac{1}{EE_1'}\cdot WW_2'\frac{1}{EE_2'}\right)-(RR_1'\cdot RR_2')<2.$$

3.1 An Example with Uniform Cost and Signal Distributions

Figure 3: Equilibria in the Clear / Unclear Example



4. Extending the Basic Model: One-Sided Policies

A. Equality in Offers

- Given beliefs (π_A, π_B) and worker application standards (t_A, t_B) , it will choose hiring standards (s_A, s_B) and make an investment decision $i_j \in \{q, u\}$ for each group $j \in \{A, B\}$ to solve the following optimization problem

$$\max_{s_A, s_B, i_A, i_B} \left[\lambda_B P\left(s_B, \pi_B, i_B\right) + \lambda_A P\left(s_A, \pi_A, i_A\right)\right] \quad \text{s.t.} \quad \rho\left(s_B, \pi_B\right) = \rho\left(s_A, \pi_A\right). \tag{6}$$

Definition. *An equilibrium under affirmative action is a set of beliefs* (π_A, π_B), (δ_A, δ_B), *worker standards* (t_A, t_B) *and employer standards* (s_A, s_B) *satisfying the following conditions:*

(a) *Firm signal thresholds* (s_A, s_B) *solve problem (6), given* (π_A, π_B, t_A, t_B).[19]

(b) $t_j = t_j^*(\delta_j)$, $j \in \{A, B\}$

(c) $\pi_j = G^W(\beta_W(s_j, \delta_j))$, $j \in \{A, B\}$

(d) $\delta_j = G^E\left(\lambda_j \left[F_u^E(t_j) - F_q^E(t_j)\right]\left[\pi_j\left(1 - F_q^W(s_j)\right)\chi_q - (1 - \pi_j)\left(1 - F_u^W(s_j)\right)\chi_u\right]\right)$, $j \in \{A, B\}$

Proposition 2. *Assume that, without affirmative action, there exists an equilibrium with positive investment. Then there exists an equilibrium under affirmative action without homogeneous beliefs.*

Proposition 3. *Assume that $\phi(\theta)$ and $\tau(\psi)$ are continuous, strictly decreasing and strictly positive on $[0,1]$. Further assume that $\lambda_A \neq \lambda_B$ and that $G^E(k)$ and $G^W(c)$ are strictly increasing. Then no equilibrium with positive investment and homogeneous employer beliefs exists (with or without affirmative action).*

Proposition 4. *Assume that $\phi(\theta)$ and $\tau(\psi)$ are continuous, strictly decreasing and strictly positive on $[0,1]$. Further suppose that the A and B markets start with $\pi_A > \pi_B > 0$ and $\delta_A > \delta_B > 0$. For fixed beliefs $\{\pi_A, \pi_B, \delta_A, \delta_B\}$ and low enough δ_B and π_B, imposing affirmative action causes zero firms to invest in B amenities and zero B workers to invest.*

B. Equality in Employment

- Given beliefs (π_A, π_B) and worker application standards (t_A, t_B) , an employer will again choose hiring standards (s_A, s_B) and make investment decisions (i_A, i_B) to solve the following problem:

$$\max_{s_A, s_B, i_a, i_B} \left[\lambda_B P\left(s_B, \pi_B, i_B\right) + \lambda_A P\left(s_A, \pi_A, i_A\right)\right] \quad \text{s.t.} \quad \rho_H\left(s_B, \pi_B, i_B\right) = \rho_H\left(s_A, \pi_A, i_A\right). \tag{7}$$

Definition. *An equilibrium under an employment quota is a set of beliefs* (π_A, π_B), (δ_A, δ_B), *worker standards* (t_A, t_B) *and employer standards* $\left(s_j^{q,q}, s_j^{q,u}, s_j^{u,q}, s_j^{u,u}\right)$, $j \in \{A, B\}$ *satisfying the following conditions:*

(a) *Each firm's investment decisions* (i_A, i_B) *and thresholds* (s_A, s_B) *solve (7), given* (π_A, π_B, t_A, t_B)

(b) $t_j = t_j^* (\delta_j)$, $j \in \{A, B\}$

(c) $\pi_j = G^W \left(\overline{\beta}_W\right)$, $j \in \{A, B\}$

(d) $\delta_j = \int_0^1 G^E \left(k_j^* (k_{-j})\right) dk_{-j}$

Proposition 5. *Assume that G^E has full support on $[0, \overline{c}]$ with $\overline{c} > \omega_q$, let $\phi(\theta)$ be strictly decreasing, and define $\tilde{s}$ as the firm signal threshold such that $\phi(\tilde{s}) = 1$. If firm investment is close enough to perfectly observable, any equilibrium under an employment quota must entail homogeneous beliefs if:*

$$\eta\left(\overline{\beta}(s)\right) < \frac{\phi(s_j)}{\phi(s_j) - 1}$$

for all $s \in [0, \tilde{s})$ where $\eta(c) = \frac{d[c \cdot G(c)]}{dc}$ and $\overline{\beta}(s) = \left[F_u^W(s) - F_q^W(s)\right]\omega_q$.

Proposition 6. *Assume that* $\phi(\theta)$ *and* $\tau(\psi)$ *are continuous, strictly decreasing and strictly positive on* $[0,1]$*. Further suppose that the A and B markets start with* $\pi_A > \pi_B > 0$ *and* $\delta_A > \delta_B > 0$*. For low enough* δ_B *and* π_B*, imposing an employment quota lowers employment of A workers. Furthermore, there exists an open set of parameters such that the policy leads to zero investment by B workers.*

C. Wage and Employments Subsidies

D. An “Impossibility” Result

Proposition 7. *Suppose that we seek to move to an equilibrium* $\{s^*, t^*, \pi^*, \delta^*\}$ *from another point with* $s_0 > s^*$, $t_0 > t^*$, $\pi_0 < \pi^*$ *by independently setting some combination* C *of* s, t, π *and* δ. *There exist interventions that achieve this aim for any* $\{\pi_0, \delta_0\}$ *if and only if* $\{\delta, \pi\} \in C$, $\{t, \pi\} \in C$ *or* $\{s, \delta\} \in C$. *Targeting* $\{\delta, \pi\}$ *is faster than any alternative.*

5. Extending the Basic Model: Two-Sided Policies

A. Two-Sided Investment Insurance

Proposition 8. *Suppose that the government observes noisy but informative signals, θ^g and ψ^g, of worker and firm investment respectively. For any initial beliefs, there exist incentive payments ω^g and χ^g conditional on these signals that immediately ensure that $\pi_A = \pi_B$, $\delta_A = \delta_B$, $s_A = s_B$ and $t_A = t_B$. If and only if $\lambda_A \neq \lambda_B$, a non-zero permanent investment subsidy is required to maintain $\pi_A = \pi_B$.*

If the government sets $s_B^g = s_A$, the fraction of B workers who invest is:

$$\pi_{B,t} = G^W\left(\beta_W\left(s^*\left(\pi_{B,t-1}\right), \delta_{B,t-1}\right) + \left[\tilde{F}_u^W\left(s_A|\theta < s^*\left(\pi_{B,t-1}\right)\right) - \tilde{F}_q^W\left(s_A|\theta < s^*\left(\pi_{B,t-1}\right)\right)\right]\omega^g\right).$$

Since this is achieved immediately, the actual cost of the worker payments are as follows.

$$\delta\left[1 - \tilde{F}_q\left(s_A|\theta < s_A\right)\right]\omega^g + (1-\delta)\left[1 - \tilde{F}_u\left(s_A|\theta < s_A\right)\right]\omega^g$$

B. Affirmative Action

Proposition 9. *Assume that* $\phi(\theta)$ *and* $\tau(\psi)$ *are continuous, strictly decreasing and strictly positive on* $[0,1]$. *For any* $\pi_B \in [0,1)$ *and* $\pi_A \in (0,1)$ *with* $\pi_B < \pi_A$, *there exist cost distributions* G^W *and* G^E, *a signal distribution* $F_i^W(\theta)$ *and parameters such that: (i)* π_B *and* π_A *are part of an equilibrium; and (ii) no one-sided investment subsidy can raise* π_B *to* π_A *in any finite number of periods* T, *even if combined with affirmative action.*

6. Interpreting Group Differences in the Presence of Two-Sided Statistical Discrimination

- Two-sided statistical discrimination complicates empirical analysis, since differences between groups are generically a combination of both employer and worker decision-making.

- For example, consider a setting with employer learning as in Altonji and Pierret (2001).

- Under conditions they outline, the conditional expectation for log-wages can be written as a time-varying function of the form:

$$E\left(w_t \middle| s_i, z_i, t\right) = b_{s,t} s_i + b_{z,t} z_i + H\left(t\right)$$

- An implication of this or any other model in which investments depend on beliefs or otherwise depend on race is that empirical analysis designed to detect statistical discrimination may be misleading.
- Assuming that race is an s variable – i.e., employers statistically discriminate – the linear predictor of the wage must be modified as follows:

$$E^*\left(w_t | s_i, z_i, t\right) = \left(b_{s,t} + \rho_{s,t}\right) s_i + b_{z,t} z_i + b_t t$$

- Lang and Lehmann (2012) discuss an alternative test that is robust to differing wage profiles of black and white workers.
- Let B_i be a dummy for whether a worker is black.
- As before, z_i is correlated with productivity and initially unobserved by the employer.
- Lang and Lehmann propose comparing two regressions.

$$E^* \left(w_t | s_i, z_i, t \right) = \alpha_1 + \alpha_2 B_i + \alpha_3 t + \alpha_4 B_i t + \alpha_5 z_i$$

$$E^* \left(w_t | s_i, z_i, t \right) = \beta_1 + \beta_2 B_i + \beta_3 t + \beta_4 B_i t + \beta_5 z_i + \beta_6 z_i t$$

- If black workers are lower productivity on average and employers statistically discriminate, we would therefore expect in the following $\gamma_4 < 0$ and $\gamma_2 > 0$ in the following auxiliary regression.

$$E^*\left(z_t t \middle| s_i, z_i, t\right) = \gamma_1 + \gamma_2 B_i + \gamma_3 t + \gamma_4 B_i t + \gamma_5 z_i$$

A. Detecting Employer Discrimination

- If workers are paid their marginal product, the wage paid by a firm to a worker with ability a_i at a firm with can k_j^F be shown to be as follows

$$\ln w_i = \left(\frac{\sigma^2_{a,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right)\ln a_i + \left(\frac{\sigma^2_{\varepsilon,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right)\mu_{a,j} + \ln\gamma + (1-\gamma)\ln k_{j,F} + \frac{1}{2}\left(\frac{\sigma^2_{\varepsilon,j}\sigma^2_{a,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right) + \left(\frac{\sigma^2_{a,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right)\ln\varepsilon_i$$

Assumption. *For a worker of long enough tenure at her previous employer, her past wage exactly reflects her ability at a new firm.*

Nonetheless, this assumption allows us to write the wage offered to worker i as a particularly simple function of her wage at her previous firm, group-specific fixed effects for the source and destination firms, and an error term:

$$\ln(w_i) = \beta_j \ln\left(w_i^{OLD}\right) + \alpha_{j,f^{OLD}} + \alpha_{j,f^{NEW}} + \nu_i \qquad (8)$$

where:

$$\alpha_{j,f^{NEW}} = \left(\frac{\sigma^2_{\varepsilon,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_a}\right)\mu_{a,j} + (1-\gamma)\,k_{j,F^{NEW}} + \frac{1}{2}\left(\frac{\sigma^2_{\varepsilon,j}\sigma^2_{a,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right)$$

$$\alpha_{j,f^{OLD}} = -(1-\gamma)\ln k_{j,F^{OLD}}$$

$$\nu_i = \left(\frac{\sigma^2_{a,j}}{\sigma^2_{\varepsilon,j}+\sigma^2_{a,j}}\right)\ln \varepsilon_i$$

and β_j is the elasticity of the wage with respect to individual ability:

Proposition 10. *Assume that ability at the new and old firms are correlated:* $\ln a_i = c_j + \rho \ln a_i^{OLD} + \ln \eta_i$ *where* $0 < \rho \leq 1$. *Conditional on firm fixed effects, assume that* $\ln \varepsilon_i$, $\ln \varepsilon_i^{OLD}$ *and* $\ln \eta_i$ *are uncorrelated with* $\ln w_i^{OLD}$ *and that a worker's past employer has more information than her new employer:* $\sigma^2_{\varepsilon,j} \geq \sigma^2_{\varepsilon,j,OLD}$. *Then the difference in coefficients from regression (8) is* $\hat{\Gamma}$, *where*

$$\hat{\Gamma} = \rho \left[\left(\frac{\sigma^2_{\varepsilon,W,OLD} + \sigma^2_{a,W}}{\sigma^2_{\varepsilon,W} + \sigma^2_{a,W}} \right) - \left(\frac{\sigma^2_{\varepsilon,B,OLD} + \sigma^2_{a,B}}{\sigma^2_{\varepsilon,B} + \sigma^2_{a,B}} \right) \right]$$

$$= \rho \left[\underbrace{\left(\frac{\sigma^2_{a,W}}{\sigma^2_{\varepsilon,W} + \sigma^2_{a,W}} \right) - \left(\frac{\sigma^2_{a,B}}{\sigma^2_{\varepsilon,B} + \sigma^2_{a,B}} \right)}_{\Gamma \text{ (true difference in returns)}} + \left(\frac{\sigma^2_{\varepsilon,W,OLD}}{\sigma^2_{\varepsilon,W} + \sigma^2_{a,W}} \right) - \left(\frac{\sigma^2_{\varepsilon,B,OLD}}{\sigma^2_{\varepsilon,B} + \sigma^2_{a,B}} \right) \right].$$

7. Conclusion