Online Appendix to

Discrimination as Retaliation

by Till Wicker

1 Extensions to Theory in Section 2

1.1 Paternalistic Discrimination (Buchmann et al., 2024)

To allow for paternalistic discrimination, the utility function of the employer also contains a fraction α_{eg} of the expected on-the-job welfare of the worker, where employers differ in whether they internalize their perception of the worker's perception of welfare, or internalize their own perception of workers' welfare. In particular, the utility function of the manager becomes:

$$\max_{L_{A,t},L_{B,t}} \underbrace{Y(L_{A,t},\theta_A,L_{B,t},\theta_B) - \sum_{g\in\{A,B\}} L_{g,t}w_g}_{\text{Firm Profit}} - \underbrace{\sum_{g\in\{A,B\}} L_{g,t}f\Big(d_g,F(\chi_{g,t})\Big)}_{\text{Non-Pecuniary Costs}} - \underbrace{\sum_{g\in\{A,B\}} L_{g,t}\alpha_{g,t}\mathcal{W}_{g,t}}_{\text{Other-regarding utility}}$$

where $\mathcal{W}_{g,t}$ is the manager's perception of the workers' perception of welfare, which is defined with respect to the outside option. The worker's welfare consists of their wage, and disutility of working ($\mathcal{W} = \mathbb{E}_i[w_g - u_g(c)]$). Paternalistic employers internalize their own perceptions of the worker's welfare.

1.2 Experience-based Discrimination (Lepage, 2024)

The employer's utility function remains as specified:

$$\max_{L_{A,t},L_{B,t}} Y(L_{A,t},\theta_A,L_{B,t},\theta_B) - \sum_{g\in\{A,B\}} L_{g,t}w_g - \sum_{g\in\{A,B\}} L_{g,t}f(d_g,F(\chi_{g,t}))$$

However, we now incorporate dynamic belief updating based on accumulated experiences with each group. In particular, at $t = 0$, employers have prior beliefs about group g's

productivity: $\hat{\theta}_{g,0} \sim N(\hat{\mu}_{g,0}, 1/\hat{\tau}_{g,0})$.

After each hiring decision, employers observe realized productivity $\theta_{g,i}$ for each hired worker i from group g. Following Bayesian updating combined with experience-based learning:

$$\hat{\mu}_{g,t+1} = \alpha_\mu \hat{\mu}_{g,t} + (1-\alpha_\mu)\left[\beta_g(\chi_{g,t}) \cdot \overline{\theta}_{g,obs,t} + (1-\beta_g(\chi_{g,t})) \cdot \hat{\mu}_{g,t}\right]$$
$$\hat{\tau}_{g,t+1} = \alpha_\tau \hat{\tau}_{g,t} + (1-\alpha_\tau)\left[\frac{H_{g,t}}{\text{var}(\theta_{g,obs,t})}\right]$$

where $\alpha_\mu, \alpha_\tau \in [0,1]$ are experience weights (higher values place more weight on past beliefs), $\beta_g(\chi_{g,t}) \in [0,1]$ is the experience-dependent learning rate from new observations, $H_{g,t}$ is the cumulative number of workers hired from group g up to time t, and $\overline{\theta}_{g,obs,t} = \frac{1}{H_{g,t}} \sum_{i=1}^{H_{g,t}} \theta_{g,i}$ is the sample mean of observed productivity.

Lepage (2024) illustrates that the learning rate itself depends on past experiences:

$$\beta_g(\chi_{g,t}) = \beta_0 \cdot \exp(-\gamma \cdot F(\chi_{g,t})) \quad \text{where } \gamma > 0, \beta_0 \in (0,1]$$

This specification captures the psychological mechanism whereby negative experiences with respect to the productivity of hired workers make employers less receptive to contradictory information. As such, past experiences can have an effect on current discrimination through two channels:

1. Learning about group-level productivity, as a result of past experiences (Experience-based discrimination, Lepage (2024))

2. Endogenously updating non-pecuniary costs of hiring workers from a specific group (Retaliatory discrimination)

1.3 Inaccurate Statistical Discrimination

The employer's utility function remains as specified in Equation (1):

$$\max_{L_{A,t}, L_{B,t}} Y(L_{A,t}, \theta_A, L_{B,t}, \theta_B) - \sum_{g \in \{A,B\}} L_{g,t} w_g - \sum_{g \in \{A,B\}} L_{g,t} f(d_g, F(\chi_{g,t}))$$

Following Bohren et al. (2025a), we now explicitly distinguish between true and subjective productivity distributions:

True Productivity Distribution: Worker productivity for group g is drawn from $\theta_g \sim N(\mu_g, 1/\tau_g)$ with true signal precision η_g.

Subjective Beliefs: Employers hold potentially inaccurate subjective beliefs $\hat{\psi} \equiv (\hat{\mu}_g, \hat{\tau}_g, \hat{\eta}_g)$ about group g's productivity distribution and signal precision, where:

$$\hat{\theta}_g \sim N(\hat{\mu}_g, 1/\hat{\tau}_g) \tag{1}$$

$$\text{Subjective signal precision:} \quad \hat{\eta}_g \geq 0, \quad \hat{\eta}_g \neq \eta_g \tag{2}$$

Inaccurate Statistical Discrimination occurs when $(\hat{\mu}_g, \hat{\tau}_g, \hat{\eta}_g) \neq (\mu_g, \tau_g, \eta_g)$ for some group g.

Following Bohren et al. (2025a), employers make hiring decisions based on their subjective posterior beliefs. After observing worker's group identity g and signal s, the employer forms a posterior belief using Bayes' rule with subjective distributions:

$$\hat{\mu}_{g,t}(s) = \frac{\hat{\tau}_g \hat{\mu}_g + \hat{\eta}_g s}{\hat{\tau}_g + \hat{\eta}_g} \tag{3}$$

1.4 Systemic Discrimination

Systemic discrimination captures how discrimination in other decisions indirectly contributes to disparities by affecting relevant attributes for a given decision, which in turn generates disparities in outcomes. This extension demonstrates how retaliatory discrimination at the focal node can coexist with systemic discrimination arising from other nodes in the decision system.

Following Bohren et al. (2025b), we embed the retaliatory discrimination model within a broader system of interconnected decision nodes. The system consists of a set of nodes $N \equiv \{1, ..., N\} \cup \{n^*\}$, where n^* represents the focal hiring node from our baseline model. Each non-focal node $n = 1, ..., N$ represents a decision task where:

- Worker i has productivity $\theta\_i^n \in \Theta^n$ for task n
- An employer observes the worker's group G_i and signal $S_i^n \in \mathcal{S}^n$

- The employer selects action $A_i^n \in \mathcal{A}^n$ according to action rule $A^n(G_i, S_i^n)$. The action is to either hire the worker, or not.

At the focal node n^* (our baseline hiring decision):

- Worker productivity is $\theta_i^* \in \Theta^*$
- Signal is $S_i^* \in \mathcal{S}^*$
- Action is $A_i^* \in \mathcal{A}^*$

As discussed in Bohren et al. (2025b), actions at other nodes can affect productivity and signals at the focal node. For example, S_i^* may include performance evaluations from other nodes, or focal-node productivity θ_i^* may depend on training received at node n.

Incorporating this within the retaliatory discrimination model, the employer's utility function at the focal node becomes:

$$\max_{L_{A,t}, L_{B,t}} Y(L_{A,t}, \theta_A, L_{B,t}, \theta_B) - \sum_{g \in \{A,B\}} L_{g,t} w_g - \sum_{g \in \{A,B\}} L_{g,t} f(d_g, F(\chi_{g,t}))$$

However, the signal S_i^* and/or productivity Y_i^* now depend on actions at other nodes:

$$S_i^* = S^*(G_i, A_i^1, A_i^2, ..., A_i^N, \xi_i)\theta_i^* \qquad = \theta^*(G_i, A_i^1, A_i^2, ..., A_i^N, \zeta_i) \tag{4}$$

where ξ_i and ζ_i represent other factors affecting signals and productivity.

As such, retaliatory discrimination at one node can have consequences for future nodes, and hence lead to more, systematized discrimination.

References

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2025a). Inaccurate Statistical Discrimination: An Identification Problem. *The Review of Economics and Statistics*, pages 1–16.

Bohren, J. A., Hull, P., and Imas, A. (2025b). Systemic Discrimination: Theory and Measurement. *The Quarterly Journal of Economics*.

Lepage, L.-P. (2024). Experience-Based Discrimination. *American Economic Journal: Applied Economics*, 16(4):288–321.