# Fairness in Selection Problems with Strategic Candidates

Vitalii Emelianov, Nicolas Gast, Patrick Loiseau

Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LIG

## Abstract

To better understand discriminations and the effect of affirmative actions in selection problems (e.g., college admission or hiring), a recent line of research proposed a model based on *differential variance*. This model assumes that the decision-maker has a noisy estimate of each candidate's quality where the difference in the noise variances between different demographic groups is a key factor to explain discrimination. The literature on differential variance, however, does not consider the strategic behavior of candidates who can react to the selection procedure to improve their outcome, which is well-known to happen in many domains.

In this paper, we study how the strategic aspect affects fairness in selection problems. We propose to model selection problems with strategic candidates as a contest game: A population of rational candidates compete by *choosing* an effort level to increase their quality. We characterize the (unique) equilibrium of this game in the different parameters' regimes, both when the decision-maker is unconstrained and when they are constrained to respect the fairness notion of demographic parity.

## Motivation: Selection with Differential Variance

quality $W$ → estimate of quality:

$$W \xrightarrow{H} \hat{W} = W + \varepsilon \cdot \sigma_H$$
$$W \xrightarrow{L} \hat{W} = W + \varepsilon \cdot \sigma_L$$

$H$ – high-noise, $L$ – low-noise

**Prior Work [Garg et al. (FAccT'21), Emelianov et al. (EC'20)]:**

- differential variance leads to discrimination
- demographic parity can improve utility of selection

**But candidates can be strategic!**

- Hardt et al., "Strategic Classification", ITCS'16.
- Kleinberg et al.,"How Do Classifiers Induce Agents to Invest Effort Strategically?", EC'19.

## Our Model of Selection with Strategic Candidates

quality (with effort $m$) → estimate of quality:

$$W \sim \mathcal{N}(\underset{\text{effort}}{m}, \eta^2) \xrightarrow{H} \hat{W} = W + \varepsilon \cdot \sigma_H$$
$$\xrightarrow{L} \hat{W} = W + \varepsilon \cdot \sigma_L$$

Each candidate $i$ attains a **individual utility** of making an effort $m$ given selection threshold $\theta$

$$u_i(\underset{\text{effort}}{m_i}, \underset{\text{threshold}}{\theta}) = \underset{\text{reward}}{S} \cdot \mathbf{P}\left(\underset{\text{expected quality}}{\mathbf{E}(W_i|\hat{W}_i)} \geq \theta\right) - \underset{\text{cost}}{C_{G_i} m_i^2}$$

We study the **Nash equilibrium**:

Decision-maker selects the best $\alpha \in (0,1)$ candidates based on $\mathbb{E}(W|\hat{W})$. A pair of effort distributions $\boldsymbol{\mu} = (\mu_H, \mu_L)$ is called an equilibrium if for all candidates $i$:

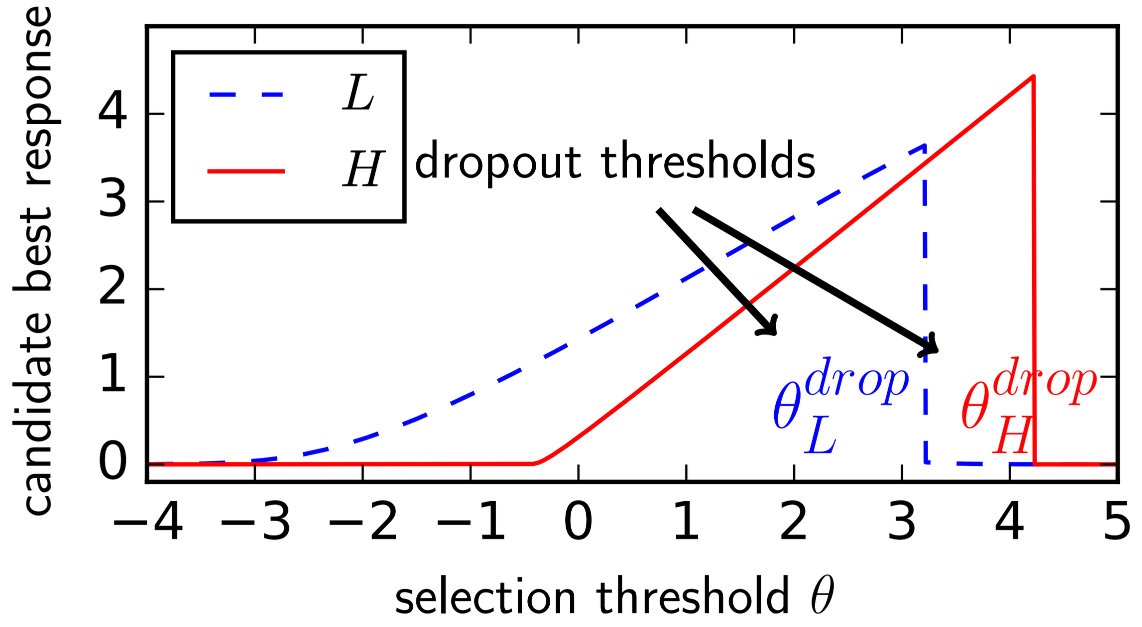
$$\forall m_i \in \text{supp}(\mu_{G_i}), m' \neq m_i : u_i(m_i, \theta(\boldsymbol{\mu})) \geq u_i(m', \theta(\boldsymbol{\mu})),$$

where $\theta(\boldsymbol{\mu}) = F_{\boldsymbol{\mu}}^{-1}(1-\alpha)$ and $F_{\boldsymbol{\mu}}$ is the CDF of expected qualities $\mathbb{E}(W|\hat{W})$ induced by the effort distribution $\boldsymbol{\mu}$

## Best Response Characterization

For large rewards $S$:

- If $C_H = C_L$, then the **low-noise** candidates drop out at a lower threshold than the high-noise candidates (see fig.)
- If $C_H \neq C_L$, then the **cost-disadvantaged** candidates drop out at a lower threshold




## Discrimination at Equilibrium

For any selection rate $\alpha \in (0,1)$ and large rewards $S$, at equilibrium:

- If $C_H = C_L$, then the **high-noise** candidates make greater effort and are overrepresented

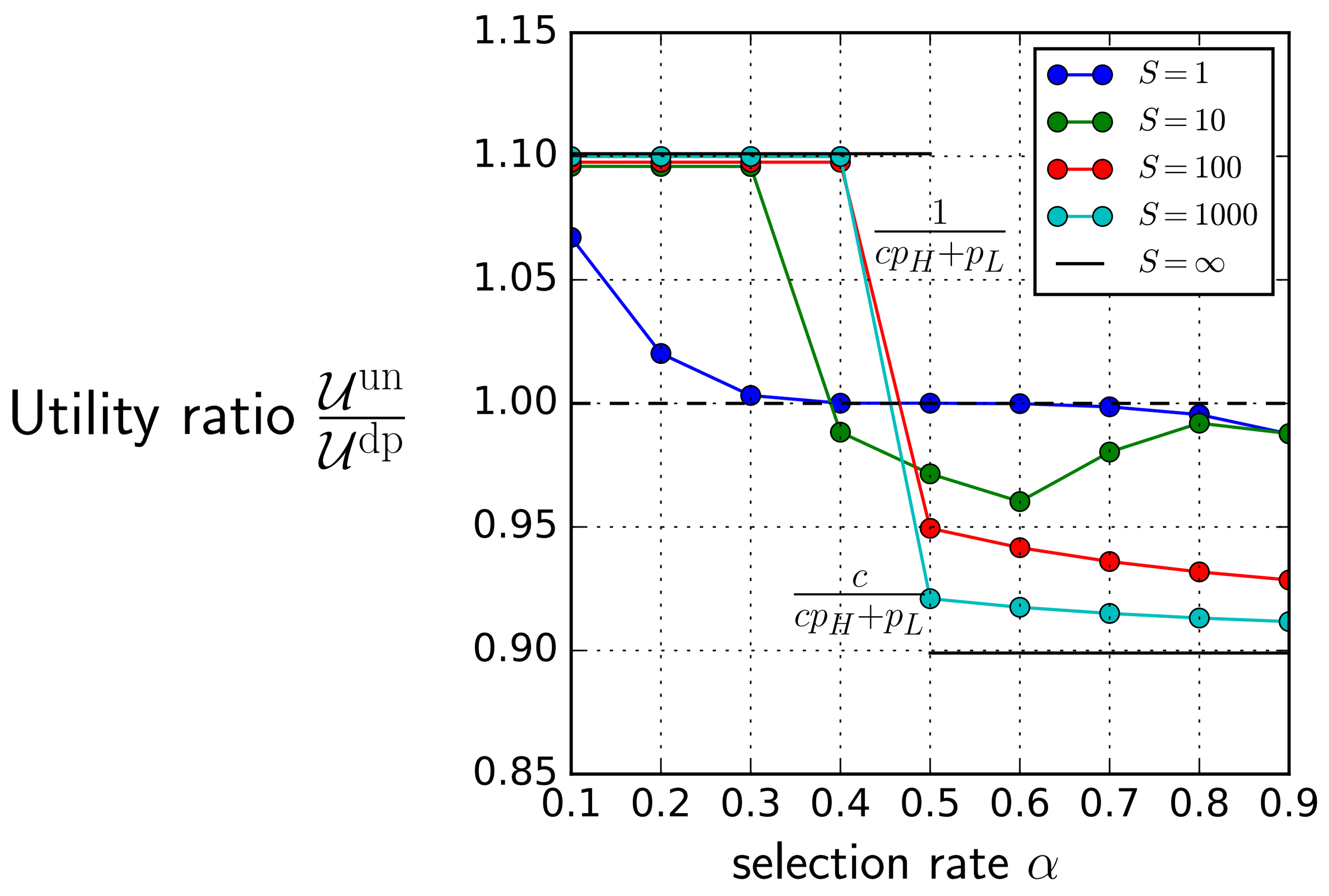
$$\lim_{S\to\infty} \frac{\bar{\mu}_L^{\text{un}}}{\bar{\mu}_H^{\text{un}}} = \lim_{S\to\infty} \frac{\bar{r}_L^{\text{un}}}{\bar{r}_H^{\text{un}}} = \begin{cases} 0 & \text{if } \alpha \geq p_H \\ (\alpha - p_H)/p_L & \text{if } \alpha < p_H \end{cases}$$

- If $C_H \neq C_L$, then the **cost-advantaged** candidates make greater effort and are overrepresented

## Effects of Demographic Parity Mechanism

For any selection rate $\alpha \in (0,1)$ and large enough reward $S$:

- Demographic parity **incentivizes** the previously underrepresented group to make a larger effort
- Demographic parity mechanism **can lead to a better qualified** selection. For $c = \sqrt{C_H/C_L} < 1$:




## Implications of the Paper

- When the cost-of-effort is group-independent, $C_H = C_L$, the discrimination is governed by the differential variance, however, the resulting discrimination is for a different group compared to the non-strategic setting
- When the cost-of-effort is group-dependent, $C_H \neq C_L$, the differential variance has a second order effect on discrimination
- Bayesian decision-making is not optimal!