Discrimination in Promotion

Anja Prummer[1], Francesco Nava[2]

[1]Queen Mary University of London

[2]London School of Economics

Corporate Culture Defines Fit

Fitting in matters especially **at promotion stage** and in sectors such as **finance, law, consulting**:

Corporate Culture Defines Fit

Fitting in matters especially **at promotion stage** and in sectors such as **finance, law, consulting**:

- Long hours
- Dress code: clothing, hairstyles
- Language: dialects/accents (Southern accent in a NYC bank)
- Hobbies: golf, marathons vs Netflix

Corporate Culture Defines Fit

Fitting in matters especially **at promotion stage** and in sectors such as **finance, law, consulting**:

- Long hours
- Dress code: clothing, hairstyles
- Language: dialects/accents (Southern accent in a NYC bank)
- Hobbies: golf, marathons vs Netflix

→ Do employers select a culture that makes a promotion accessible/ appealing for all workers?

Corporate Culture Defines Fit

Fitting in matters especially **at promotion stage** and in sectors such as **finance, law, consulting**:

- Long hours
- Dress code: clothing, hairstyles
- Language: dialects/accents (Southern accent in a NYC bank)
- Hobbies: golf, marathons vs Netflix

→ Do employers select a culture that makes a promotion accessible/ appealing for all workers?

→ Does an employer benefit from inducing differential valuations for a promotion among his workers?

Discrimination through Culture

Workers with private value for promotion, employer knows distribution of valuations

Discrimination = Design of worker's value distributions

$\rightarrow$ work environment, organisational culture

Discrimination through Culture

Workers with private value for promotion, employer knows distribution of valuations

Discrimination = Design of worker's value distributions

$\rightarrow$ work environment, organisational culture

Employer's Maximisation Problem

$$\max_{\text{Value Distributions}} \quad \text{Sum of Worker's Effort}$$
$$\text{s.t.} \quad \text{Constraints on Distributions}$$

Constraints on Values

(1) Benchmark: no constraint

(2) Value Dispersion: culture leads to adjustment in values of one worker, does not impact values of other worker

- Adjusted distribution cannot lead to higher average valuation $\rightarrow$ encompasses SOSD
- Designed distribution is first order stochastically dominated by given distribution

(3) Value Reallocation: culture affect workers differentially

- Adjustment matches some measure over values

Key Intuition

Key trade-off: Distribution design leads to

(1) reduction in information rent as worker's value more recognisable
(2) inequalities between workers reducing competition

Key Intuition

Key trade-off: Distribution design leads to

(1) reduction in information rent as worker's value more recognisable

(2) inequalities between workers reducing competition

$\rightarrow$ **Redistribution of value increases effort**

$\rightarrow$ **Discrimination is profitable**

Related Literature

Organisational Incentives and Culture

Winter 2004; Kreps 1990; Crémer 1993, Lazear 1995; Hodgson 1996; Hermalin 1999, 2012; Gibbons + co-authors 2015, 2020

→ Inequality with substitutable effort through work culture

Mechanism Design/Information Design

Myerson 1981; Condorelli, Szentes 2019; Rösler, Szentes 2017; Bergemann, Pesendorfer 2007; Sorokin, Winter 2018; Bobkova 2019; Haghpanah, Ali, Lin, Siegel 2020

→ Mechanism design subject to ”capacity constraint”

Contests with Handicap

Lazear, Rosen 1981; Mealem, Nitzan 2016; Li, Yu 2012; Franke 2012; Calsamiglia, Franke, and Rey-Biel 2013

→ Incentive to make agents unequal

Outline

1. Model of Discrimination
2. Benchmark: No Constraints
3. Value Dispersion
4. Value Reallocation
5. Discussion

Model of Discrimination

- Employer maximises total effort of its 2 workers, A and B
- Worker i exerts effort e_i
- 2 workers compete for a promotion through effort which depends on their value
- Workers value the promotion at v_i
- Valuation is independent, private value distributed with cdf F_i on support $[\alpha_i, \omega_i] \subseteq [0, \overline{\omega}]$, $\overline{\omega} < \infty$
- Workers start with distribution $G(v) = F_A(v) = F_B(v)$
- Probability of being promoted: x_i

Payoffs

1. Worker's expected payoff

$$x_i(\mathbf{v})v_i - e_i(\mathbf{v})$$

2. Employer's expected payoff

$$\mathbb{E}[e_A] + \mathbb{E}[e_B]$$

$\rightarrow$ Agents maximize payoffs

- the worker by choosing the optimal effort given his valuation and the probability of promotion
- the employer by implementing the optimal mechanism and selecting the value distribution, subject to constraints

Employer's Optimal Mechanism (Myerson 1981)

Direct mechanism specifies:

- an effort rule $e(\mathbf{v})$ specifies effort of worker
- allocation rule $x(\mathbf{v})$ pinning down probability of promotion

$\rightarrow$ rules are incentive compatible (IC) and individually rational (IR)

Total effort in IC and IR mechanism
= expected virtual surplus if virtual value is regular, $\psi_i'(v_i) \geq 0$:

$$TE\left(F_A, F_B\right) = \mathbb{E}_{\mathbf{v}}\left[\sum_i e_i(\mathbf{v})\right] = \underbrace{\mathbb{E}_{\mathbf{v}}\left[\sum_i \psi_i(v_i) x_i(\mathbf{v})\right].}_{\text{Expected Virtual Surplus}}$$

where

Virtual Valuation $$\psi_i(v_i) = \underbrace{v_i}_{\text{Value}} - \underbrace{\frac{1 - F_i(v_i)}{f_i(v_i)}}_{\text{Information Rent}}$$

Quantile Space

In our setting: adjustment of distributions of values

$\rightarrow$ Regularity may fail

$\rightarrow$ Quantile Space

Define

Quantile	$q_i(v_i) = 1 - F_i(v_i)$
Value	$v_i(q) = F_i^{-1}(1-q)$
Virtual Value	$\phi_i(q) = \psi_i(v_i(q)) = \frac{\partial\,(v_i(q)q)}{\partial q}$
Promotion Probability	$y_i(q), y_i'(q) \leq 0$

Quantile Space

In our setting: adjustment of distributions of values
$\rightarrow$ Regularity may fail
$\rightarrow$ Quantile Space

Define

Quantile	$q_i(v_i) = 1 - F_i(v_i)$
Value	$v_i(q) = F_i^{-1}(1-q)$
Virtual Value	$\phi_i(q) = \psi_i(v_i(q)) = \frac{\partial (v_i(q)q)}{\partial q}$
Promotion Probability	$y_i(q), y_i'(q) \leq 0$

Maximisation Problem in Quantile Space

$$\max_{F_A, F_B} \quad TE(F_A, F_B) = \mathbb{E}_q[\phi_A(q)y_A(q)] + \mathbb{E}_q[\phi_B(q)y_B(q)]$$

s.t. constraints on distributions

Employer's Constraints

1. Benchmark: no constraint
2. Value Dispersion: work environment focusing on one worker
   - Adjustment cannot lead to higher average valuation

     $$\mathbb{E}_F(v) \leq \mathbb{E}_G(v)$$

     $\rightarrow$ encompasses SOSD
   - Designed distribution is first order stochastically dominated

     $$F(v) \geq G(v)$$

3. Value Reallocation: organisational culture favours one worker, disadvantages the other
   - Distributions match some measure $H(v)$ with mass 2

     $$F_A(v) + F_B(v) = H(v) = 2G(v)$$

Outline

1. Model of Discrimination
2. Benchmark: No Constraints
3. Value Dispersion
4. Value Reallocation
5. Discussion

Benchmark: Employer's Problem without Constraints

Proposition

If the employer can adjust the value distribution for both workers arbitrarily, then he assigns measure 1 to value $\overline{\omega}$ for at least one worker.

- Employer wants workers' values to be as precise as possible $\rightarrow$ atom
  - Knowing worker's value reduces information rent paid to ensure incentive compatibility.
  - If distribution is single atom $\rightarrow$ zero information rent, employer can extract all the effort a worker with value $\overline{\omega}$ is willing to exert
- Employer wants worker to exert as much effort as possible $\rightarrow$ effort increasing in value $\rightarrow$ choose highest possible value
- Influencing one distribution sufficient

Outline

1. Model of Discrimination
2. Benchmark: No Constraints
3. **Value Dispersion**
4. Value Reallocation
5. Discussion

Value Dispersion: Lower Means

Keep distribution of worker A fixed, adjust distribution of worker B

CONSTRAINT:

$$\mathbb{E}_{F_B}[v] \leq \mathbb{E}_G[v]$$

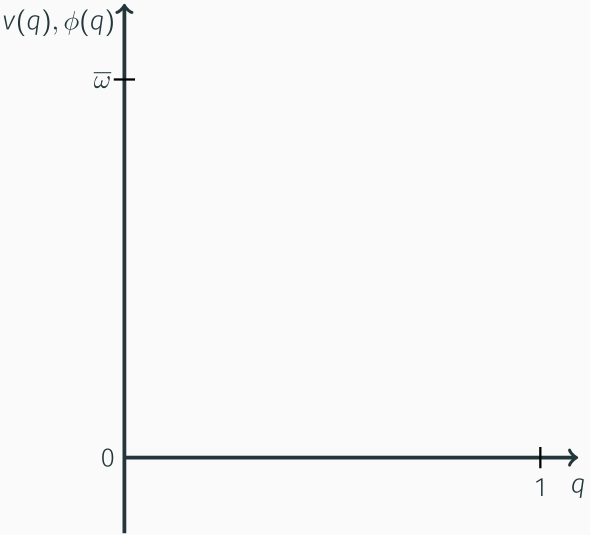
Proposition

Adjustment of B's distribution to

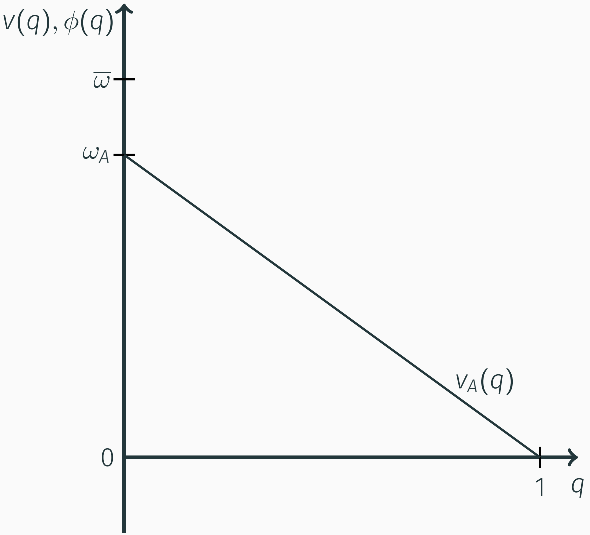
$$F^*(v) = 1 - \frac{\mathbb{E}_G[v]}{\overline{\omega}} \qquad \forall\, 0 \leq v < \overline{\omega}, \qquad F^*(\overline{\omega}) = 1$$

maximises total effort among all distributions F with $\mathbb{E}_F[v] \leq \mathbb{E}_G[v]$.

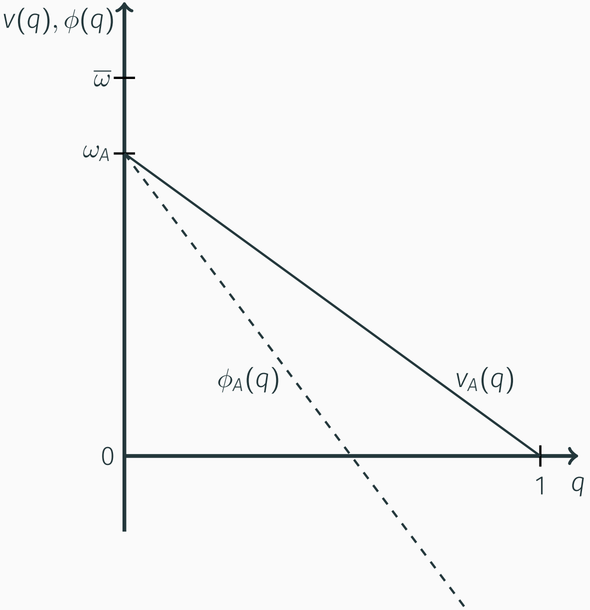
B's Value Adjustment



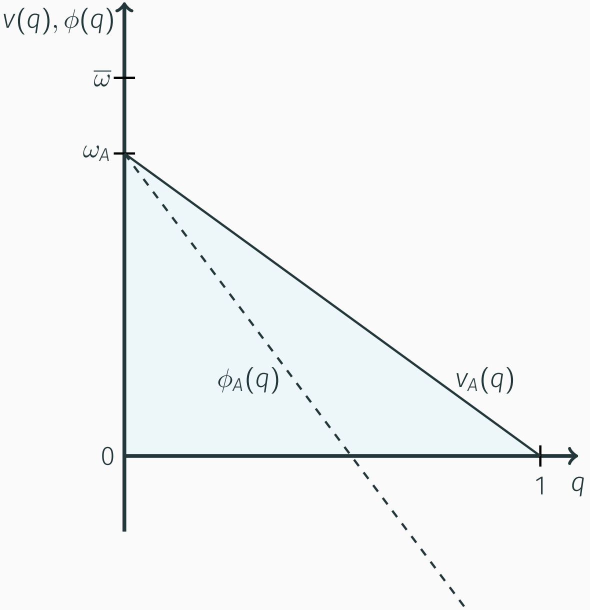
B's Value Adjustment



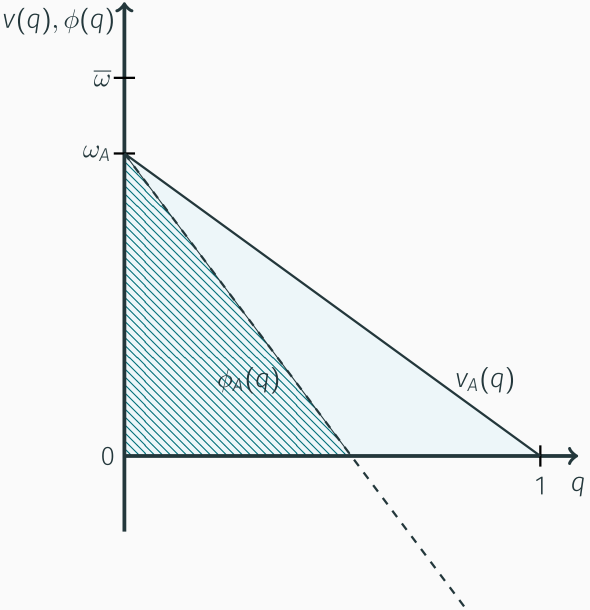
B’s Value Adjustment



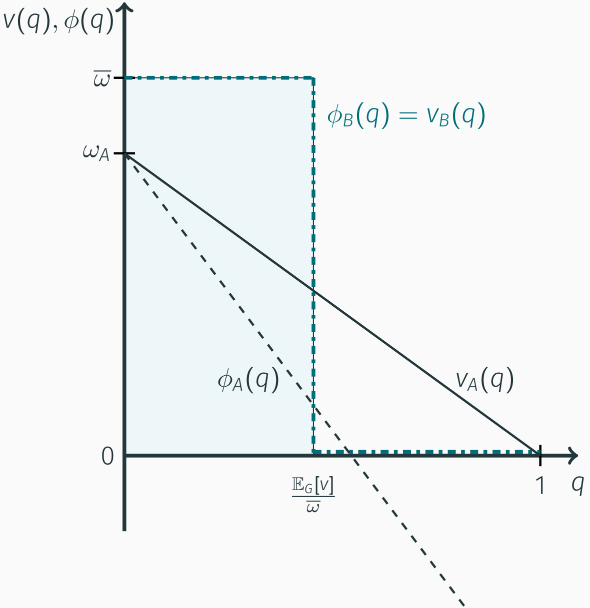
B's Value Adjustment



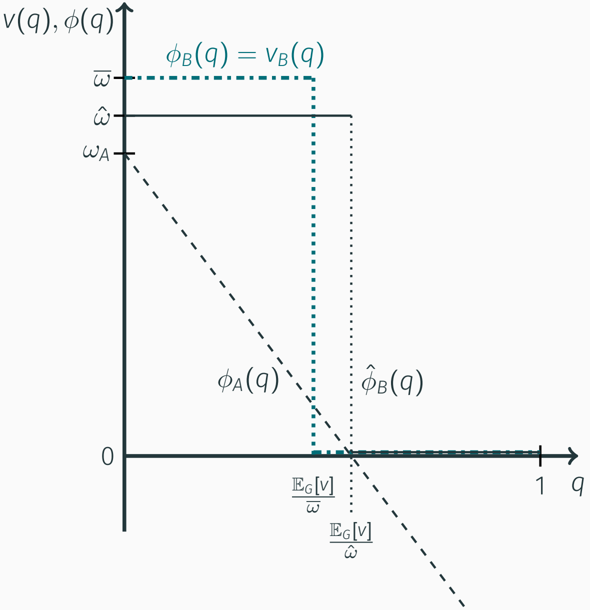
B's Value Adjustment



B's Value Adjustment



B's Value Adjustment



B's Adjustment: Intuition

Adjustment of distribution has 2 effects

- Employer knows B's value if positive, value irrelevant if zero
  $\rightarrow$ reduces information rent to zero
  Virtual value for adjusted distribution

$$\phi_B(q) = \begin{cases} \overline{\omega} & \text{if } q < \frac{\mathbb{E}_G[v]}{\overline{\omega}} \\ 0 & \text{if } q > \frac{\mathbb{E}_G[v]}{\overline{\omega}} \end{cases}$$

- Employer maximizes the probability of obtaining the promotion for worker A
  $\rightarrow$ choose atom at $\overline{\omega}$ as minimizes probability of non-zero value for B
  $\rightarrow$ induces A to exert higher effort

SOSD

CONSTRAINT

$$\int_0^v G(t)dt \leq \int_0^v F_B(t)dt \text{ for all } v \in [0, \overline{\omega}].$$

→ distributions are second order stochastically dominated by initial distribution

Corollary

Among all distributions that are second order stochastically dominated by G, F^* maximises total effort.

- Employer selects "riskiest" value distribution for worker B
- Employer has a worker A with smooth value distribution

→ A as a safe option, B as risky option

Reducing Expected Valuation

Costs to influence the distribution $\rightarrow$ adjustment still worthwhile

Corollary

For any $E_G[v] > m \geq \mathbb{E}_G[\max\{\psi(v), 0\}]$, *the distribution*

$$F^+(v) = \begin{cases} 1 & \text{if } v = \overline{\omega} \\ 1 - \frac{m}{\overline{\omega}} & \text{if } v < \overline{\omega} \end{cases}$$

yields higher total effort compared to no adjustment.

- Reduction in expected value, $E_G[v] - m$, compensated by reduction in information rent, $E_G[v] - E_G[\max\{\psi(v), 0\}]$
- Adjustment is optimal as long as reduction in expected value is lower than reduction in information rent

Value Dispersion: Lower Means for Both Workers

Employer adjusts the distribution of both workers

Corollary

Assume $\mathbb{E}_{F_i}[v] \leq \mathbb{E}_G[v], \forall i \in \{A, B\}$. *Total effort is maximised by setting for worker i,*

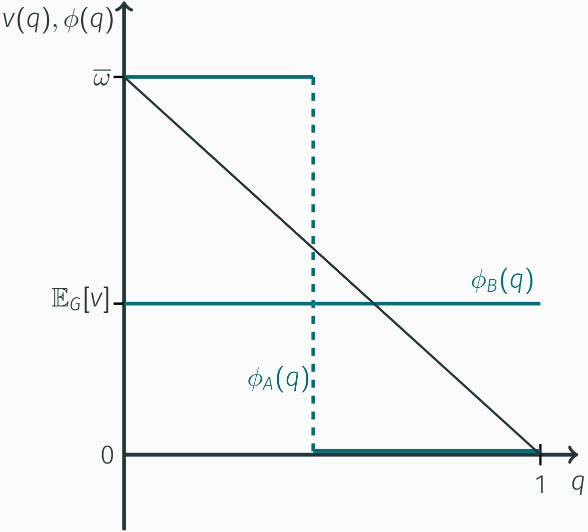
$$F_i^*(v) = \begin{cases} 1 & \text{if } v = \overline{\omega} \\ 1 - \frac{\mathbb{E}_G[v]}{\overline{\omega}} & \text{if } v < \overline{\omega} \end{cases}$$

and for worker j either (i) $F_j^*(v) = F_i^*(v)$ *or (ii)*

$$F_j^*(v) = \begin{cases} 1 & \text{if } v \geq \mathbb{E}_G[v] \\ 0 & \text{if } v < \mathbb{E}_G[v] \end{cases}$$

- Employer reduces information rent to zero for both workers
- Distributions of both workers can be the same or maximally different

Adjustment A and B: A Picture



Destroying Value: FOSD

Worker B faces certain environment, reducing his value such that new distribution is first order stochastically dominated

CONSTRAINT:

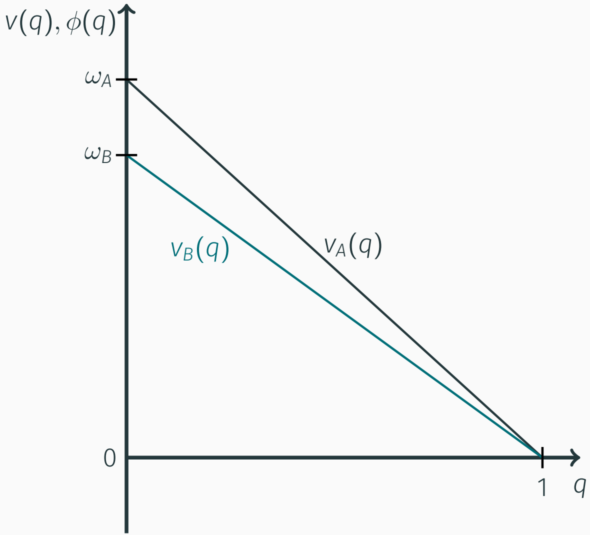
$$F(v) \geq G(v)$$

Proposition

Total effort is maximised by having no discrimination and setting $F_B(v) = F_A(v) = G(v)$ *among all distributions* $F(v) \geq G(v)$ $\forall v$.

- Adjustment of distributions in FOSD-sense never optimal
- To reduce information rent, value has to be decreased drastically $\Rightarrow$ too much to make adjustment worthwhile

FOSD in Quantile – Value Space: A dominates B



FOSD: Comparison of Effort

1. A suboptimal allocation rule lowers expected virtual surplus

$$\mathbb{E}\Big[\phi(q)y(q)\Big] \geq \mathbb{E}\Big[\phi(q)\hat{y}(q)\Big]$$

2. Integration by parts yields

$$\mathbb{E}\Big[\phi(q)y(q)\Big] = \mathbb{E}\Big[v(q)q(-y'(q))\Big]$$

3. If A dominates B then, it must hold that

$$\mathbb{E}\Big[\Big(v_A(q) - v_B(q)\Big)\left(-y'_B(q)\right)q\Big] > 0$$

$\rightarrow$ any distribution that first order stochastically dominates B yields higher surplus

Value Dispersion: An Overview

- Bi-modal distribution optimal
    1. Optimal adjustment does not destroy expected value
    2. Worker's value is recognisable, no information rent
    3. To not discourage other workers, select value such that a high value worker is least likely to occur
    4. Result encompasses second order stochastically dominated distributions
    5. If cost from adjusting distributions, still optimal to adjust (for sufficiently low costs)
- Adjustment to first order stochastically dominated distributions never optimal
  $\rightarrow$ Making value more recognisable reduces value too much

Outline

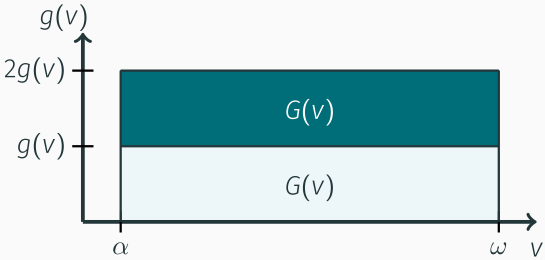
1. Model of Discrimination
2. Benchmark: No Constraints
3. Value Dispersion
4. **Value Reallocation**
5. Discussion

Value Reallocation: Match Distributions

CONSTRAINT:

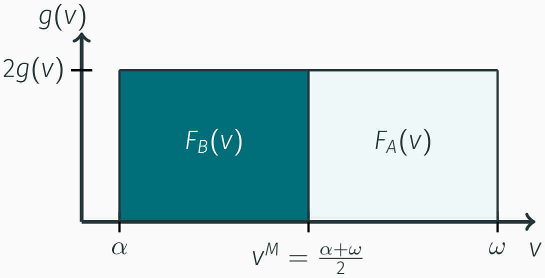
$$F_A(v) + F_B(v) = H(v) = 2G(v)$$

An example: $v \sim Uniform[\alpha, \omega]$



Fixed Measure: Example Adjusted

Adjustment: both distributions are as distinct as possible:
$v_B \sim Uniform[\alpha, \frac{\alpha+\omega}{2}]$ and $v_A \sim Uniform[\frac{\alpha+\omega}{2}, \omega]$



Fixed Measure: Beyond Example

Proposition

Reallocating value such that

$$F_B(v) = H(v) \qquad \text{for } v \in [\alpha, v^M]$$
$$F_A(v) = H(v) - 1 \qquad \text{for } v \in [v^M, \omega]$$

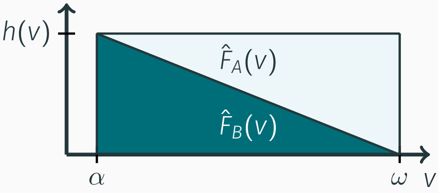
maximises total effort.

Key Insight: Minimise information rent, $\frac{1-F(v)}{f(v)}$

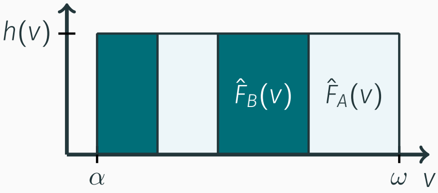
- Information rent inversely proportional to $F(v)$: high for lower values
- Maximal discrimination assigns highest mass to low values $\rightarrow$ yields maximal total effort as minimises information rent **for low values**

Fixed Measure: Proof Sketch

Compare maximal discrimination to



(a) Splitting Densities



(b) Disjoint Support

Discussion

Fixed Measure: Splitting Densities

- Denote by $a(v)$ share of density $h(v)$ assigned to A, $1 - a(v)$ share assigned to B under $\hat{F}_A$, $\hat{F}_B$
- Define auxiliary distributions $\overline{F}_A$, $\overline{F}_B$ such that

$$\text{for } v \geq v^M \qquad \overline{x}_A(v) = \max\{\hat{x}_A(v), \hat{x}_B(v)\}$$
$$\text{for } v < v^M \qquad \overline{x}_B(v) = \max\{\hat{x}_A(v), \hat{x}_B(v)\}$$

- Total effort under $\hat{F}_A$, $\hat{F}_B$ equals total effort under $\overline{F}_A$, $\overline{F}_B$
- Compare total effort with allocation rule $\overline{x}_A(v), \overline{x}_B(v)$

$$\underbrace{\mathbb{E}_v[\psi_A(v)\overline{x}_A(v)] + \mathbb{E}_v[\psi_B(v)\overline{x}_B(v)]}_{\text{Virtual Value: } A,B\text{, Allocation: } \overline{A},\overline{B}}$$
$$> \qquad \underbrace{\mathbb{E}_v[\overline{\psi}_A(v)\overline{x}_A(v)] + \mathbb{E}_v[\overline{\psi}_B(v)\overline{x}_B(v)]}_{\text{Virtual Value: } \overline{A},\overline{B}\text{, Allocation: } \overline{A},\overline{B} = \text{ Virtual Value: } \hat{A},\hat{B}\text{, Allocation: } \hat{A},\hat{B}}$$

Fixed Measure: Splitting Densities 1

Comparison of effort split into (i) $v \geq v^M$ and (ii) $v \leq v^M$

$$\int_{v^M}^{\omega} \overbrace{\left[\psi_A(v)\overline{x}_A(v) - \overline{\psi}_A(v)\overline{a}(v)\overline{x}_A(v) - \overline{\psi}_B(v)(1-\overline{a}(v))\overline{x}_B(v)\right]}^{\text{Difference in virtual values, weighted by allocation probabilities: } v>v^M} h(v)dv$$

$$+ \int_{\alpha}^{v^M} \underbrace{\left[\psi_B(v)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{a}(v)\overline{x}_A(v) - \overline{\psi}_B(v)(1-\overline{a}(v))\overline{x}_B(v)\right]}_{\text{Difference in virtual values, weighted by allocation probabilities: } v<v^M} h(v)dv > 0$$

$\rightarrow$ replace $\overline{x}_B(v)$ by $\overline{x}_A(v)$ for $v > v^M$
$\rightarrow$ replace $\overline{x}_A(v)$ by $\overline{x}_B(v)$ for $v < v^M$

Fixed Measure: Splitting Densities 2

Expression in brackets simplifies to

$$\left[\psi_i(v) - \overline{\psi}_A(v)\overline{a}(v) - \overline{\psi}_B(v)(1 - \overline{a}(v))\right] \overline{x}_i(v)$$

→ comparison of virtual values
→ simplifies to comparison of information rents

1. For $i = A$ difference is zero at each value
   → any two distributions lead to the same information rent for values above the median
2. For $i = B$ difference is one at each value
   → maximal discrimination saves on information rent for values below the median

→ Information rent high at low values, inversely proportional to $F(v)$

Maximal discrimination assigns highest mass possible to low values
→ yields maximal total effort as minimises information rent for low values

Outline

Corporate Culture in Reality

1. bi-modal distribution: consistent with divisive culture in law, banking and consultancies culture with long hours
   $\rightarrow$ loved by few, disliked by most
2. culture disadvantages e.g. women more: women face expectation to spend time with family and focus on work
   $\rightarrow$ lose-lose situation for women (Padavic, Ely, Reid 2020)

$\rightarrow$ gender, race, background determine fit

Relation to Other Sources of Discrimination

1. Taste-based Discrimination (Becker 1957):
   individuals dislike those who are different from them
   $\rightarrow$ competed away, different individuals not hired
2. Statistical Discrimination (Phelps 1972, Arrow 1973):
   exogenous or endogenous differences between groups lead to distinct outcomes of groups
   $\rightarrow$ multiplicity of equilibria, discrimination if coordination failure
   but: workers are hired with less information compared to promotion stage, statistical discrimination should be less important at later stages (Bohren Imas Rosenberg 2019; Altonji Pierret 2001)

Conclusion

- Employer benefits from redistributing workers' valuation for promotion, but not from destruction (FOSD)
- Employer aims for workers' valuation to be as recognisable as possible while maximising competition between workers
- Creating more recognisable workers reduces information rent and gain in information rent generally outweighs loss in competition

→ impact of corporate culture on workers
→ novel source of discrimination
→ model of designing value distributions

Discrimination is profitable

Fixed Measure: Disjoint Support

- At least one distribution must have disjoint support
- Analyse problem in value-quantile-space as solution boils down to comparison of quantiles
- Define auxiliary allocation probability, keeping total effort constant
- Difference in quantiles is some constant
- Possible to generate reduction in information rent in $v > v^M$, but at cost of increase in information rent for lower v
- Reduction in information rent for high values is never as high as that for low values

Fixed Measure: Disjoint Support

- At least one distribution must have disjoint support
- Analyse problem in value-quantile-space as solution boils down to comparison of quantiles
- Define auxiliary allocation probability, keeping total effort constant
- Difference in quantiles is some constant
- Possible to generate reduction in information rent in $v > v^M$, but at cost of increase in information rent for lower v
- Reduction in information rent for high values is never as high as that for low values

Maximal discrimination assigns highest mass possible to low values
$\rightarrow$ yields maximal total effort as minimises information rent for low values