Assessing Group Fairness with Social Welfare Optimization
Part 2

John Hooker
Carnegie Mellon University

Joint work with

Violet (Xinying) Chen
Stevens Institute of Technology

Derek Leben
Carnegie Mellon University

ISAIM 2024

Alpha Fairness

- Recall the **alpha fairness** SWF:

$$W_\alpha(\boldsymbol{u}) = \begin{cases} \frac{1}{1-\alpha}\sum_i u_i^{1-\alpha} & \text{for } \alpha \geq 0,\ \alpha \neq 1 \\ \sum_i \log(u_i) & \text{for } \alpha = 1 \end{cases}$$

where u_i is the utility allocated to individual i

  - **Utilitarian** when $\alpha = 0$, **maximin** when $\alpha \to \infty$
  - **Proportional fairness** (Nash bargaining solution) when $\alpha = 1$

- To achieve alpha fairness:

Maximize $W_\alpha(\boldsymbol{u})$ subject to resource constraints.

Alpha Fairness

- Alpha fair selection

Let $x_i = 1$ if individual i is selected, 0 otherwise.
Then $u_i = a_i x_i + b_i$, where $a_i =$ **selection benefit**
$b_i =$ base utility .

Now

$$W_\alpha(\boldsymbol{u}) = \begin{cases} \dfrac{1}{1-\alpha}\displaystyle\sum_i (a_i x_i + b_i)^{1-\alpha} & \text{for } \alpha \geq 0,\ \alpha \neq 1 \\ \displaystyle\sum_i \log(a_i x_i + b_i) & \text{for } \alpha = 1 \end{cases}$$

We want to maximize $W_\alpha(\boldsymbol{u})$ subject to $x_i \in \{0,1\}$ and

$$\sum_i x_i = m$$

← Number of individuals selected

Alpha Fairness

- A simple solution algorithm using algebraic trick

If $\alpha \neq 1$, we have

$$W_\alpha(\boldsymbol{u}) = \boxed{\frac{1}{1-\alpha}\sum_i b_i^{1-\alpha}} + \frac{1}{1-\alpha}\sum_i \left((a_i x_i + b_i)^{1-\alpha} - b_i^{1-\alpha}\right)$$

Constant term

Alpha Fairness

- A simple solution algorithm using algebraic trick

If $\alpha \neq 1$, we have

$$W_\alpha(\boldsymbol{u}) = \frac{1}{1-\alpha}\sum_i b_i^{1-\alpha} + \boxed{\frac{1}{1-\alpha}\sum_i \Big((a_i x_i + b_i)^{1-\alpha} - b_i^{1-\alpha}\Big)}$$

So we can maximize

$$\sum_{i|x_i=1} \frac{1}{1-\alpha}\Big((a_i + b_i)^{1-\alpha} - b_i^{1-\alpha}\Big) = \sum_{i|x_i=1} \boxed{\Delta_i(\alpha)}$$

Welfare differential
of individual *i*

Alpha Fairness

- A simple solution algorithm using algebraic trick

If $\alpha \neq 1$, we have

$$W_\alpha(\boldsymbol{u}) = \frac{1}{1-\alpha}\sum_i b_i^{1-\alpha} + \boxed{\frac{1}{1-\alpha}\sum_i \Big((a_i x_i + b_i)^{1-\alpha} - b_i^{1-\alpha}\Big)}$$

So we can maximize

$$\sum_{i|x_i=1} \frac{1}{1-\alpha}\Big((a_i + b_i)^{1-\alpha} - b_i^{1-\alpha}\Big) = \sum_{i|x_i=1} \boxed{\Delta_i(\alpha)}$$

Welfare differential
of individual *i*

... by selecting the m individuals with the largest welfare differentials $\Delta_i(\alpha)$.

Similarly (using logs) if $\alpha = 1$.

Alpha Fairness Example

$\alpha = 0.7$, Select 9 individuals

Majority group

a_i	$\Delta_I(0.7)$
1.5	**0.750**
1.4	**0.708**
1.3	**0.665**
1.2	**0.621**
1.1	**0.577**
1.0	**0.531**
0.9	0.484
0.8	0.436
0.7	0.387
0.6	0.336

Protected group

a_i	$\Delta_I(0.7)$
0.2	0.187
0.4	0.354
0.6	0.505
0.8	0.643
1.0	0.770

Alpha Fairness Example

α = 0.7, Select 9 individuals

Majority group

a_i	$\Delta_i(0.7)$
1.5	**0.750**
1.4	**0.708**
1.3	**0.665**
1.2	**0.621**
1.1	**0.577**
1.0	**0.531**
0.9	0.484
0.8	0.436
0.7	0.387
0.6	0.336

Protected group

a_i	$\Delta_i(0.7)$
0.2	0.187
0.4	0.354
0.6	**0.505**
0.8	**0.643**
1.0	**0.770**

9 individuals with highest welfare differentials

a_i	$\Delta_i(0.7)$
1.0	**0.770**
1.5	**0.750**
1.4	**0.708**
1.3	**0.665**
0.8	**0.643**
1.2	**0.621**
1.1	**0.577**
1.0	**0.531**
0.6	**0.505**

Alpha Fairness Example

α = 0.7, Select 9 individuals

- Alpha fairness (α = 0.7) corresponds to demographic parity.
  - 6 of 10 majority individuals selected
  - 3 of 5 protected individuals selected
  - 60% of both groups

9 individuals with highest welfare differentials

a_i	$\Delta_i(0.7)$
1.0	0.770
1.5	0.750
1.4	0.708
1.3	0.665
0.8	0.643
1.2	0.621
1.1	0.577
1.0	0.531
0.6	0.505

Alpha Fairness Example

α = 0.7, Select 9 individuals

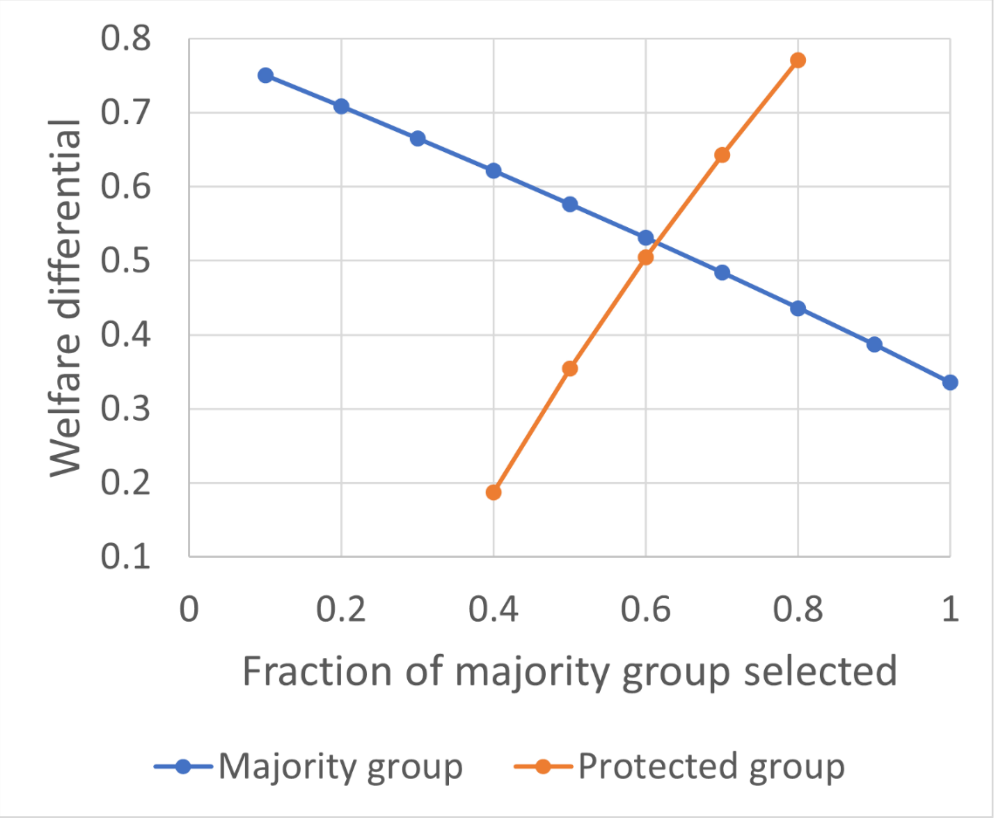
Majority group

a_i	$\Delta_I(0.7)$
1.5	**0.750**
1.4	**0.708**
1.3	**0.665**
1.2	**0.621**
1.1	**0.577**
1.0	**0.531**
0.9	0.484
0.8	0.436
0.7	0.387
0.6	0.336

Protected group

a_i	$\Delta_I(0.7)$
0.2	0.187
0.4	0.354
0.6	**0.505**
0.8	**0.643**
1.0	**0.770**

Graphical interpretation



Utility Model for 2 Groups

- We want a model that relates alpha fairness to the utility characteristics of the majority and projected groups.
  - …while reducing the number of utility parameters
  - Selection benefits uniformly distributed in each group
  - Base utility is constant in each group

Majority group

Selection benefits

$A_{\max}$ —— $A_{\min}$

Base utility = *B*

Protected group

Selection benefits

$a_{\max}$ —— $a_{\min}$

Base utility = *b*

Utility Model for 2 Groups

- We want a model that relates alpha fairness to the utility characteristics of the majority and projected groups.
  - …while reducing the number of utility parameters

Let $S =$ fraction of majority group selected
$s =$ fraction of protected group selected

Then the welfare differential of the last individual selected in the majority group is

$$\Delta_S(\alpha) = \begin{cases} \frac{1}{1-\alpha}\Big(\big((1-S)A_{\max} + SA_{\min} + B\big)^{1-\alpha} - B^{1-\alpha}\Big) & \text{if } \alpha \neq 1 \\ \log\big((1-S)A_{\max} + SA_{\min} + B\big) - \log(B) & \text{if } \alpha = 1 \end{cases}$$

and in the protected group is $\Delta'_s(\alpha)$, similarly defined.

Utility Model for 2 Groups

If β = fraction of population that is in the protected group
σ = fraction of population selected, then

$$(1-\beta)S + \beta s = \sigma,$$

which implies

$$s = s(S) = \frac{\sigma - (1-\beta)S}{\beta}$$

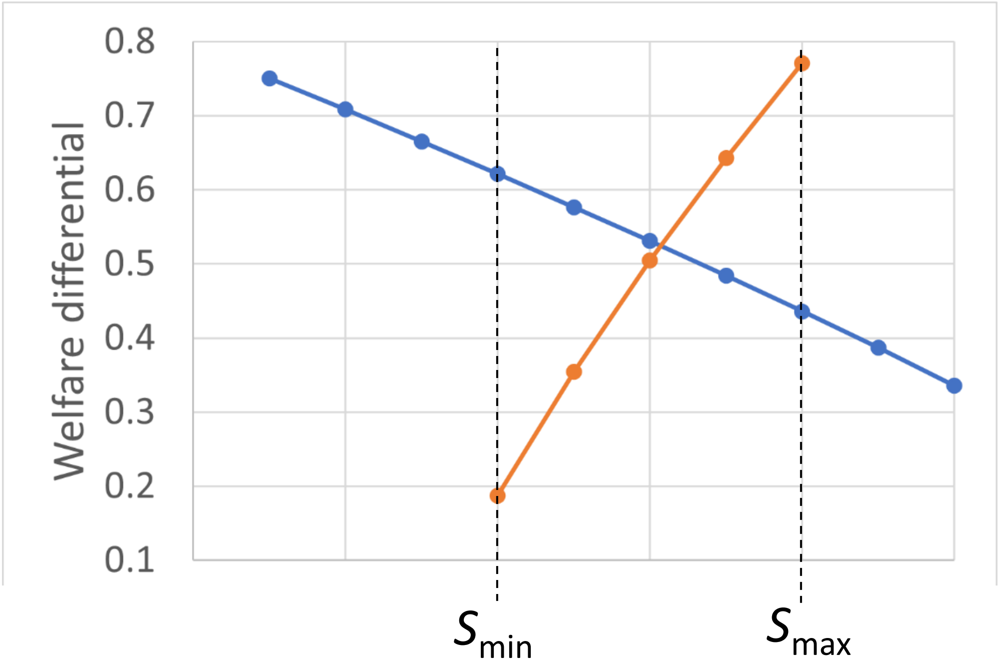
and…

Utility Model for 2 Groups

If β = fraction of population that is in the protected group
σ = fraction of population selected, then

the min and max values of S are

$$S_{\min} = \max\left\{0,\ \frac{\sigma - \beta}{1 - \beta}\right\},\quad S_{\max} = \min\left\{1,\ \frac{\sigma}{1 - \beta}\right\}$$

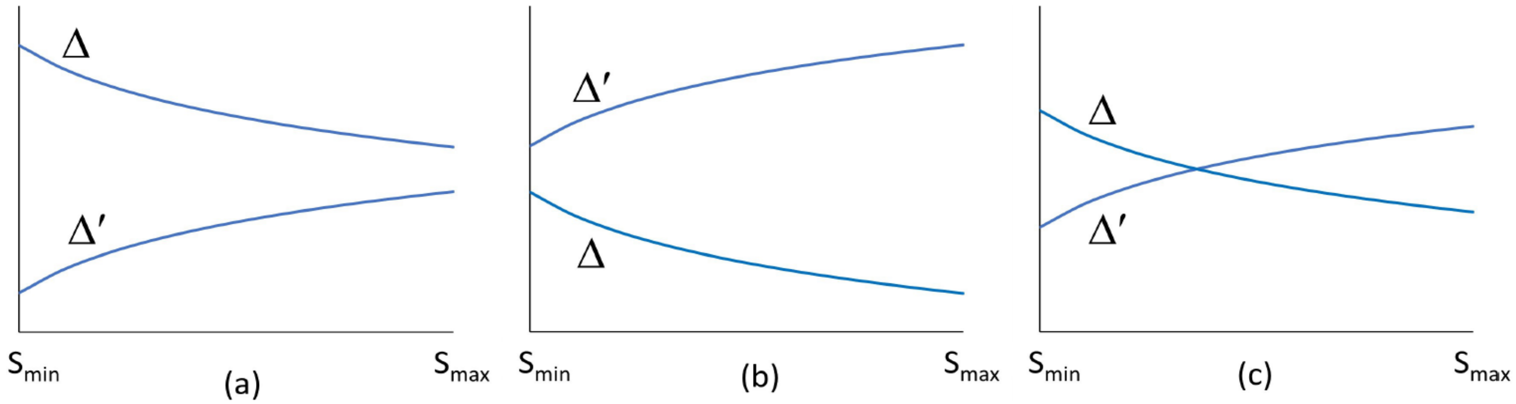


Utility Model for 2 Groups

Theorem. Selction rates (S, s) achieve alpha fairness for a given α if and only if

$$\begin{cases} (S,s) = \left(\min\left\{1, \frac{1}{1-\beta}\right\}, \frac{\sigma}{\beta}\left[1 - \min\left\{1, \frac{1-\beta}{\sigma}\right\}\right]\right) & \text{in case(a)} \\ (S,s) = \left(\frac{\sigma}{1-\beta}\left[1 - \min\left\{1, \frac{\beta}{\sigma}\right\}\right], \min\left\{1, \frac{\sigma}{\beta}\right\}\right) & \text{in case (b)} \\ \Delta_S(\alpha) = \Delta'_s(\alpha) & \text{in case (c)} \end{cases}$$

where the cases are



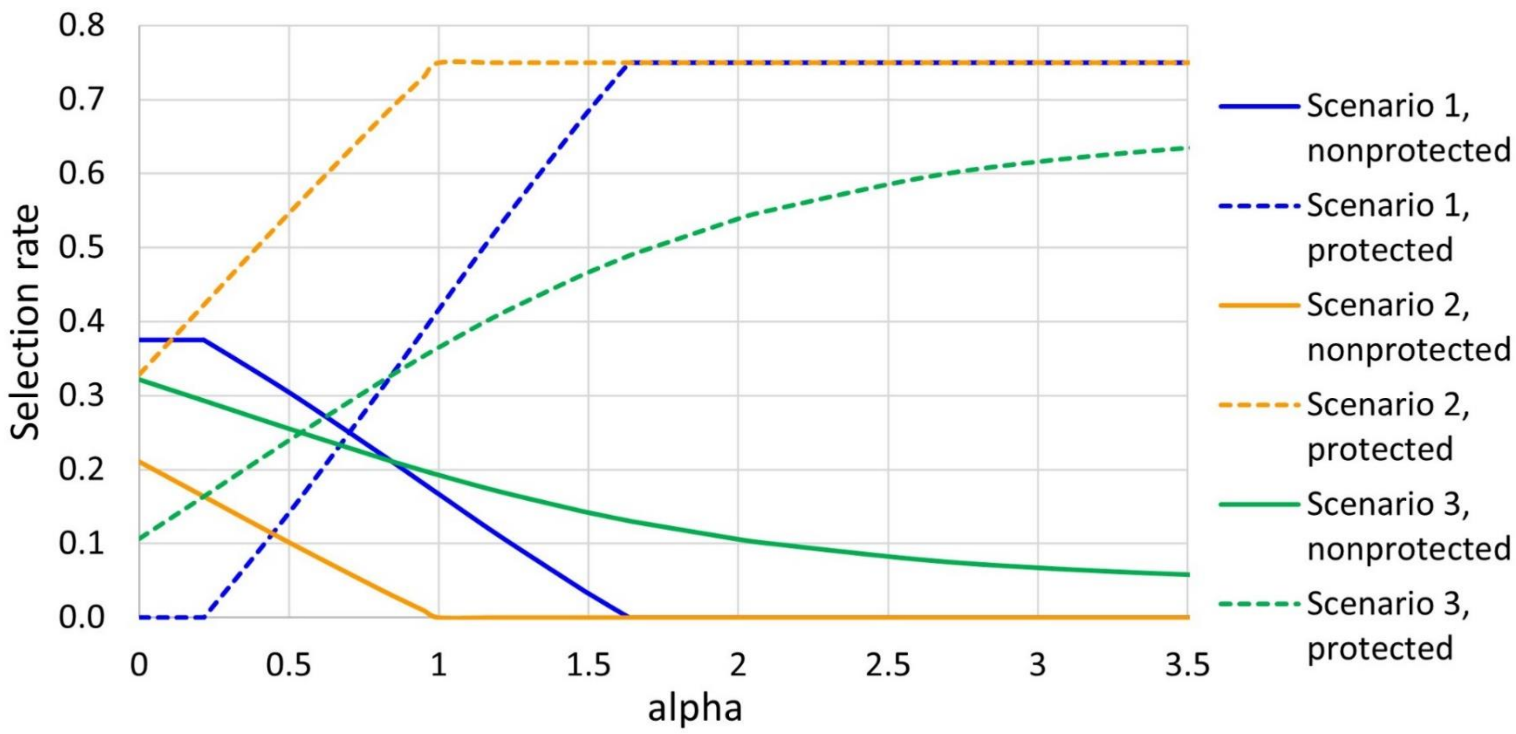
Alpha-fair Selection Rates

- Recall the 3 utility scenarios….

	Scenario 1	Scenario 2	Scenario 3
Majority group	$A_{\min}$ 0.5 — 1.5 $A_{\max}$	$A_{\min}$ 0.5 — 0.8 $A_{\max}$	$A_{\min}$ 0.5 — 1.0 $A_{\max}$
Protected group	$a_{\min}$ 0.2 — 1.0 $a_{\max}$	$a_{\min}$ 0.2 — 1.0 $a_{\max}$	$a_{\min}$ −0.5 — 1.0 $a_{\max}$
	Protected group benefits somewhat less from selection	Some protected individuals benefit more	Some protected individuals harmed by selection

Alpha-fair Selection Rates

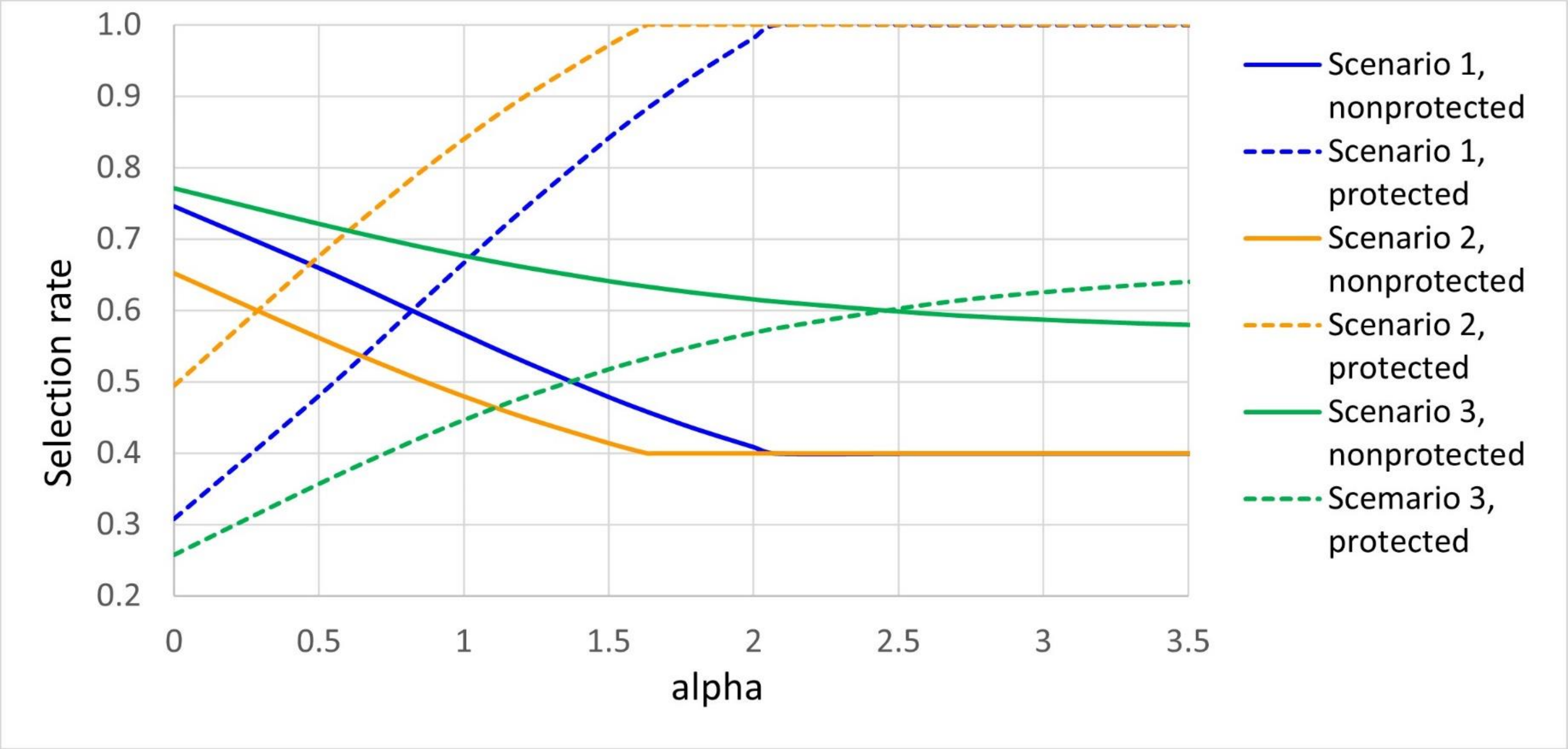
- Overall selection rate = 0.25



- Protected group has lower selection rates in Scenario 1 than in Scenario 2 due to higher utility cost of fairness in scenario 1.
- Protected group selection rate approaches 2/3 asymptotically because 1/3 of group is harmed by selection.

Alpha-fair Selection Rates

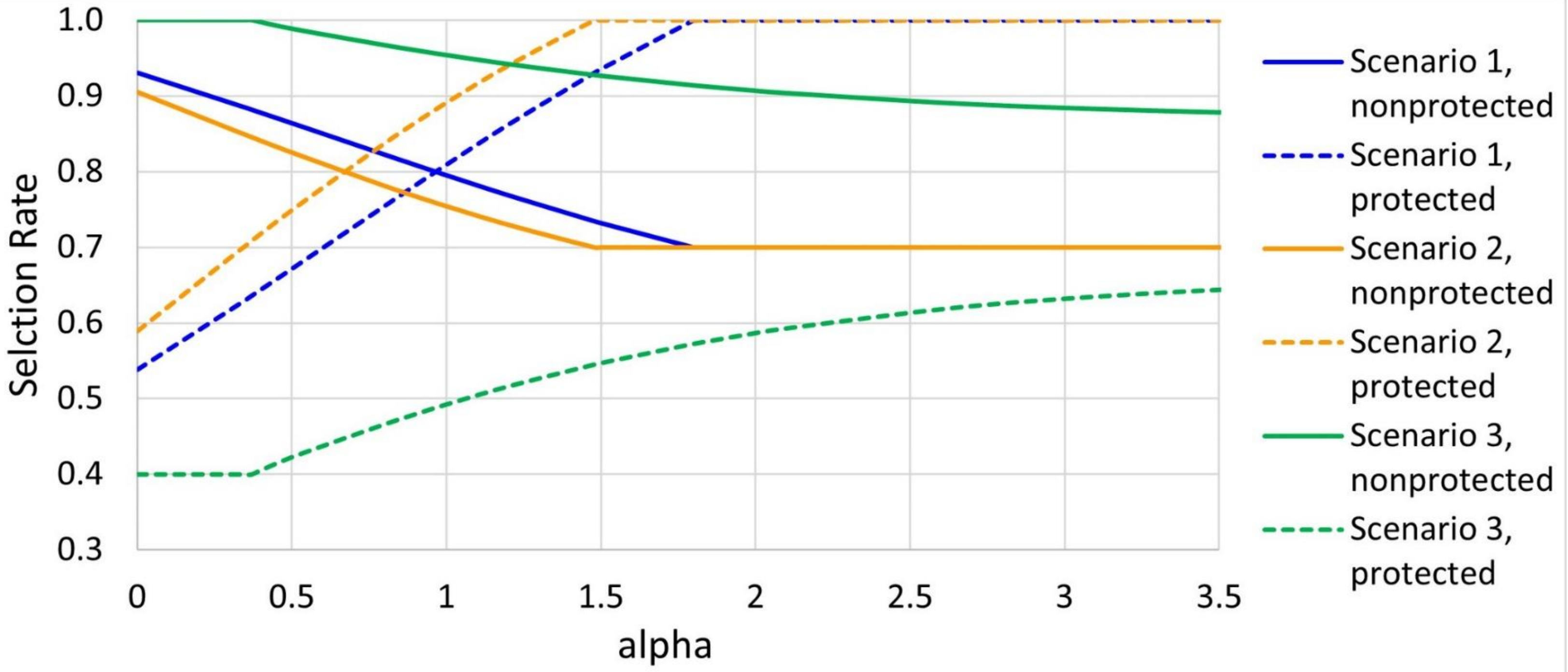
- Overall selection rate = 0.6



- Similar pattern, higher rates.

Alpha-fair Selection Rates

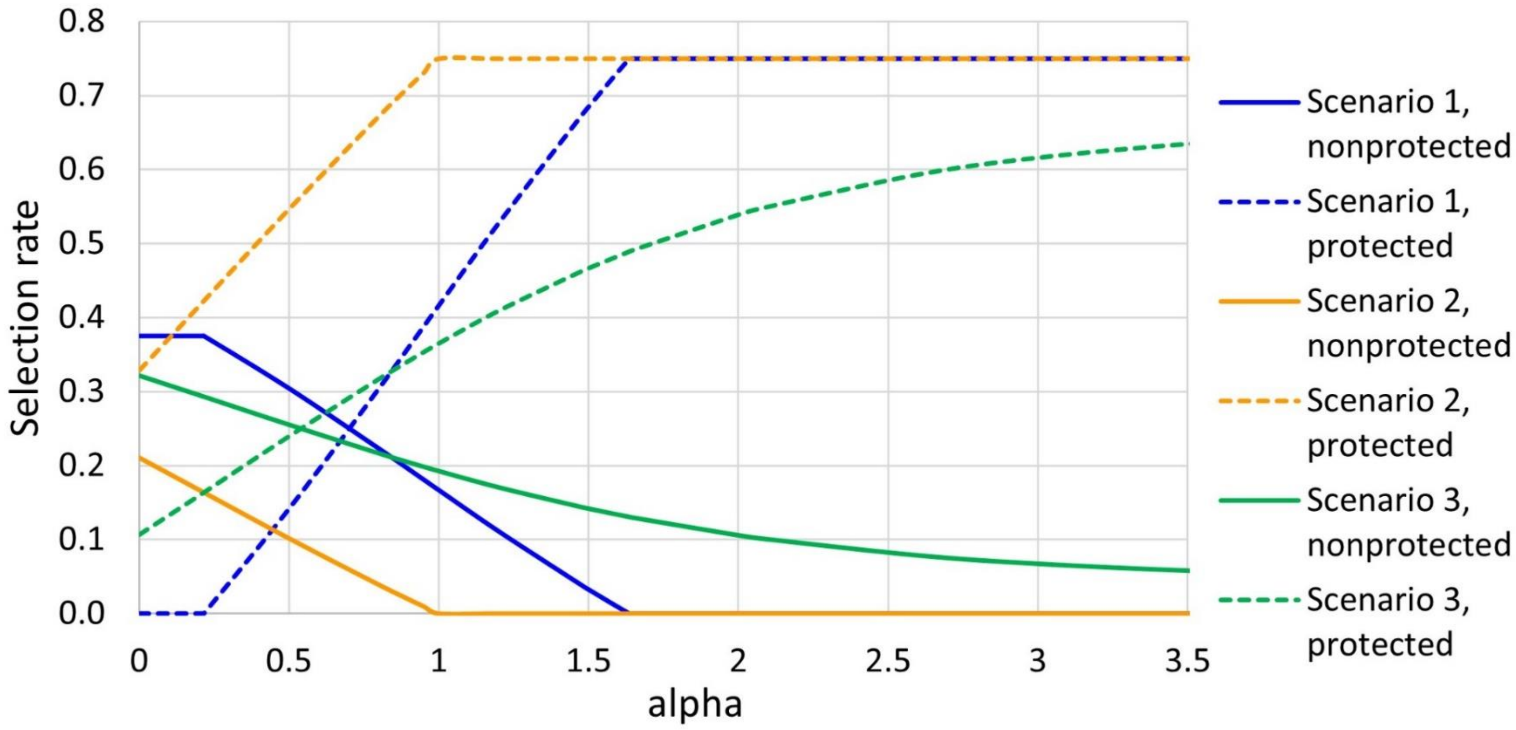
- Overall selection rate = 0.8



- Similar pattern, still higher rates.

Demographic Parity

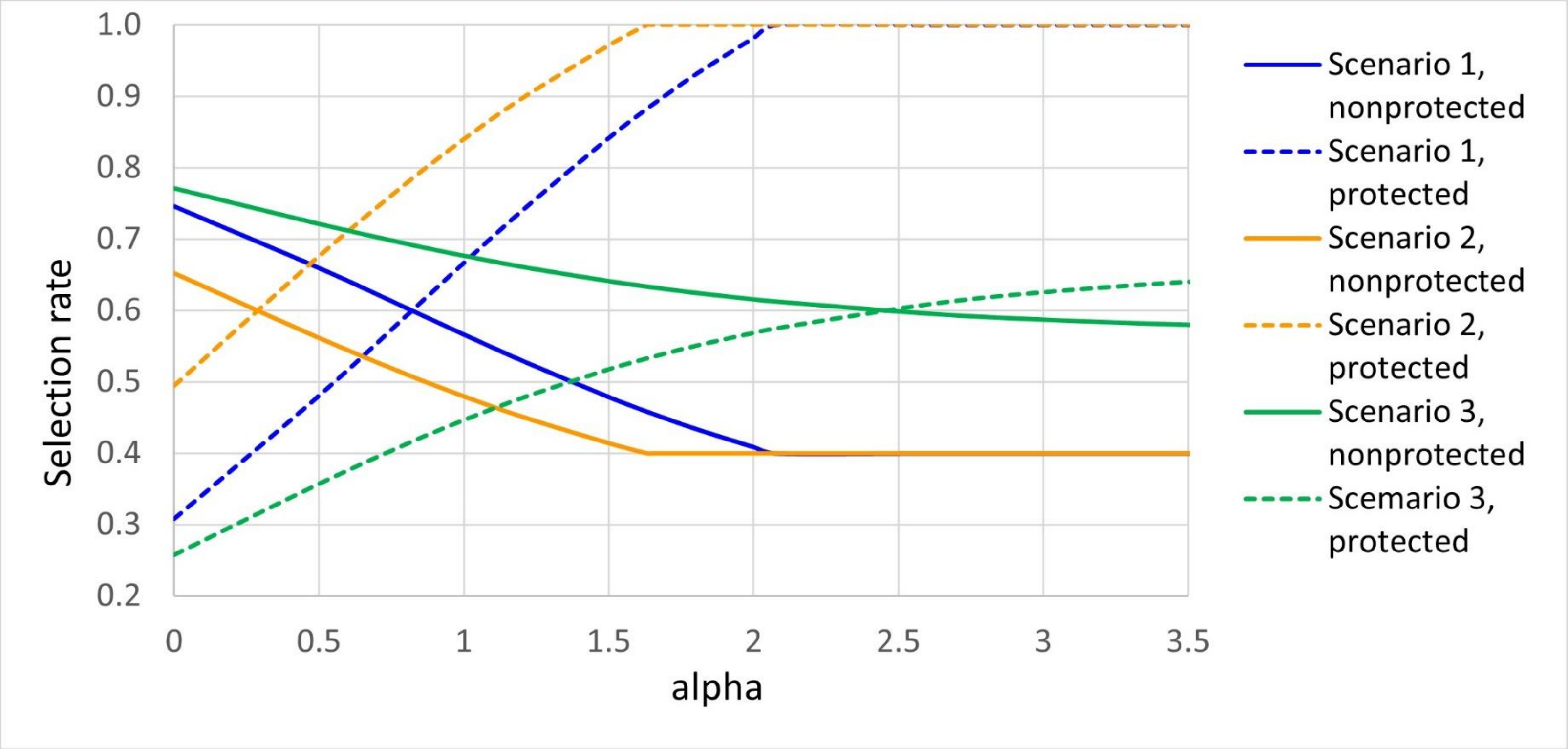
- Overall selection rate = 0.25



- Parity achieved when majority & protected curves intersect.
- Parity corresponds to relatively low degree of fairness.
- Protected group in Scenario 2 has higher rate even with $\alpha = 0$.

Demographic Parity

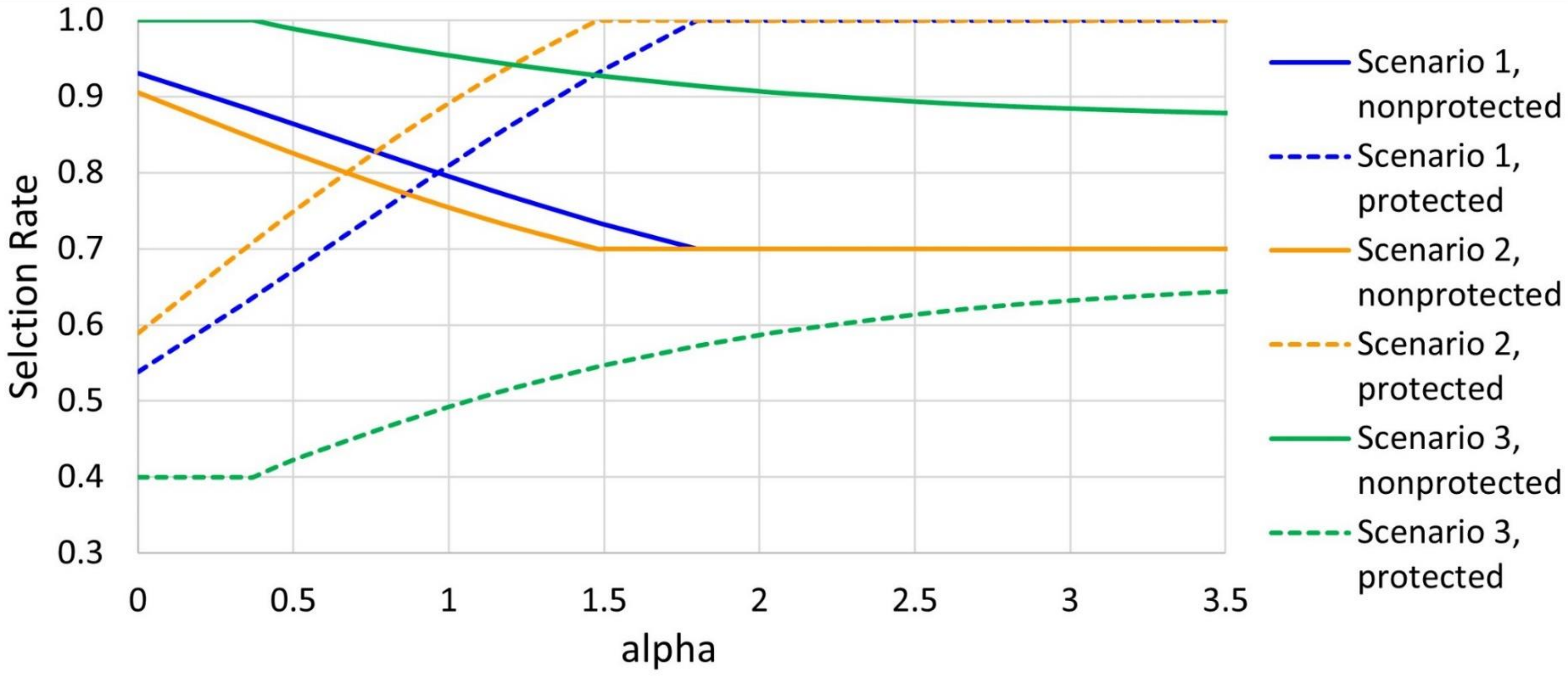
- Overall selection rate = 0.6



- Parity in Scenario 2 now requires a slight degree of fairness.
- Scenario 3 parity requires large α due to high cost of fairness.

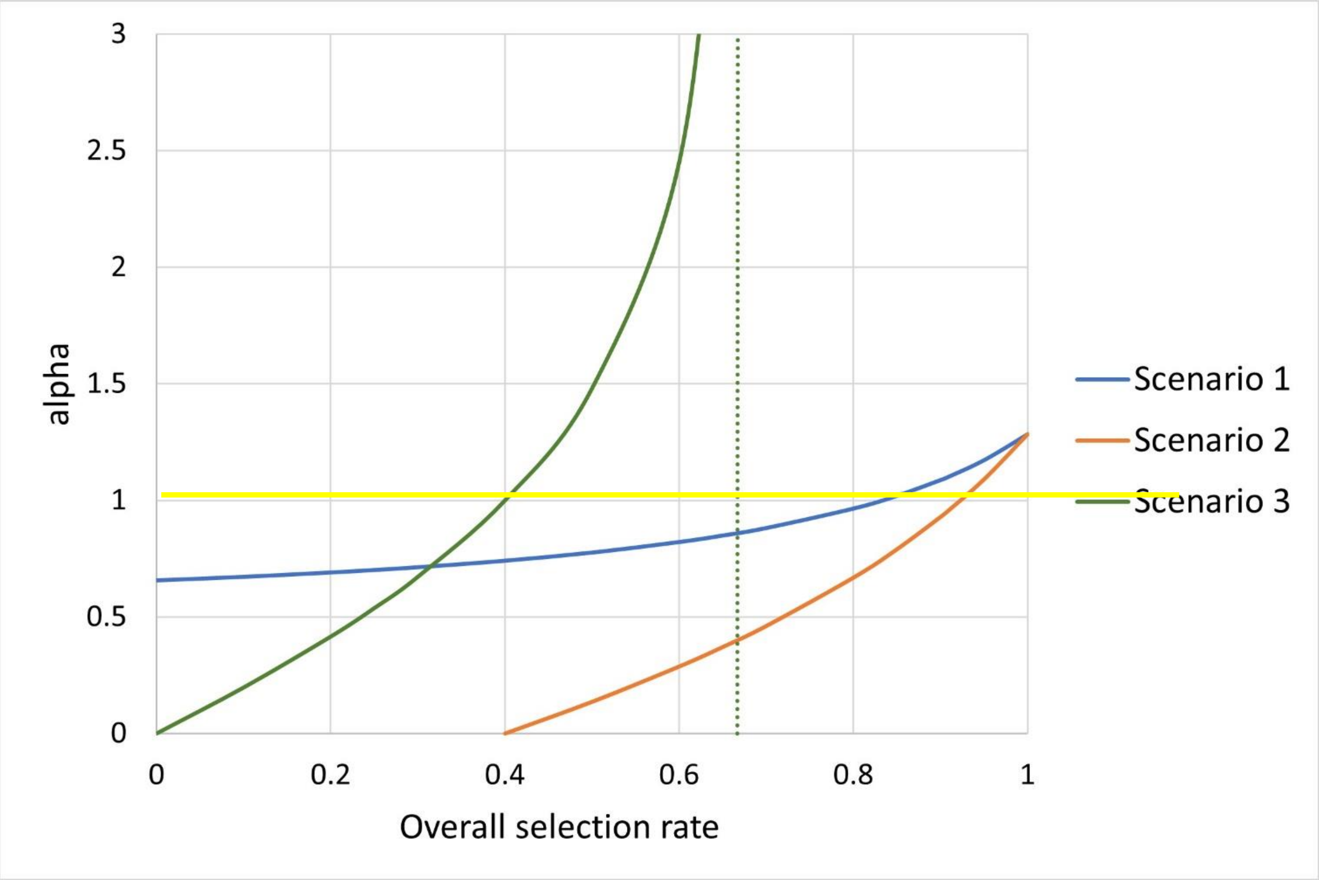
Demographic Parity

- Overall selection rate = 0.8



- Parity impossible in Scenario 3 because alpha fairness never calls for harmful selections.

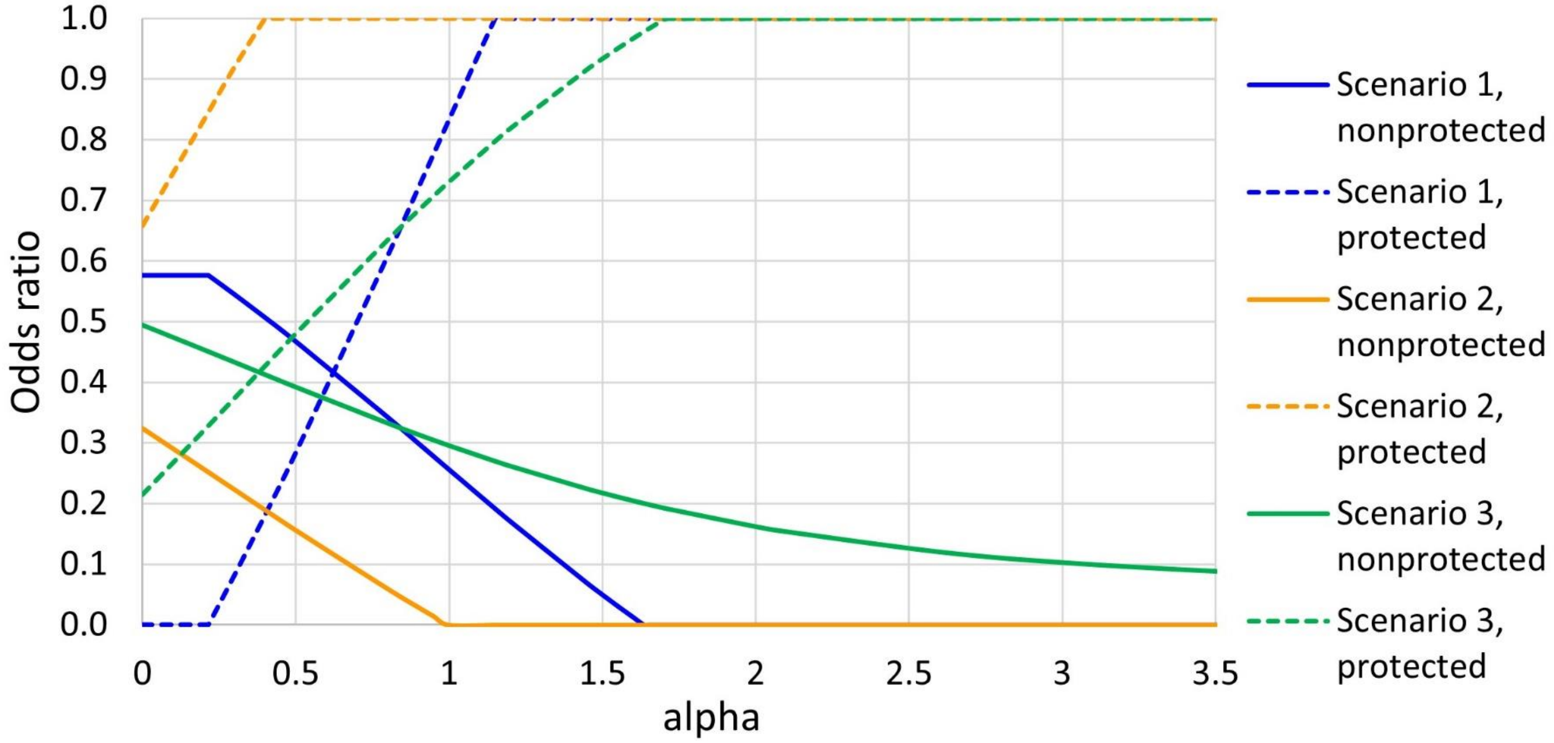
Demographic Parity



- Parity generally corresponds to **less than proportional fairness**.

Equalized Odds

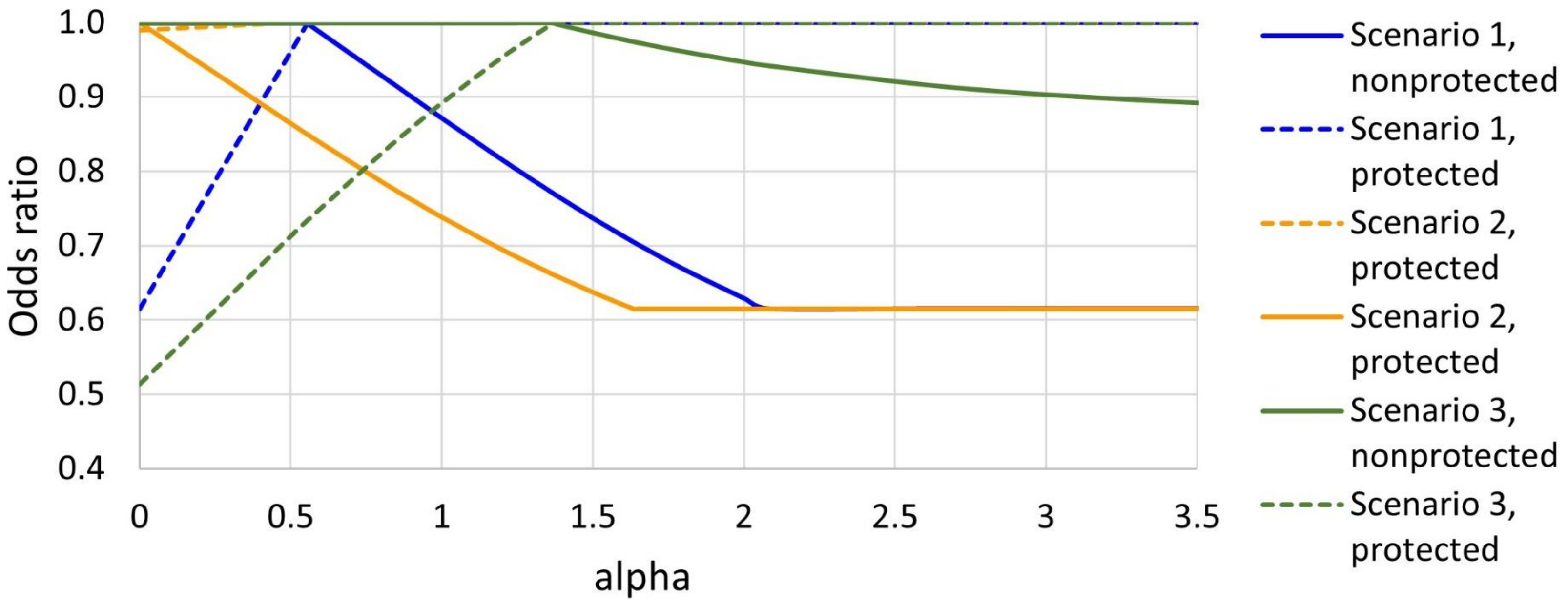
- Assume majority is 65% qualified, protected group 50% qualified.
- Overall selection rate = **0.25** < overall qualification rate of 0.6



- Even **less fair than demographic parity**.
- Sometimes viewed as easier to defend than demographic parity.

Equalized Odds

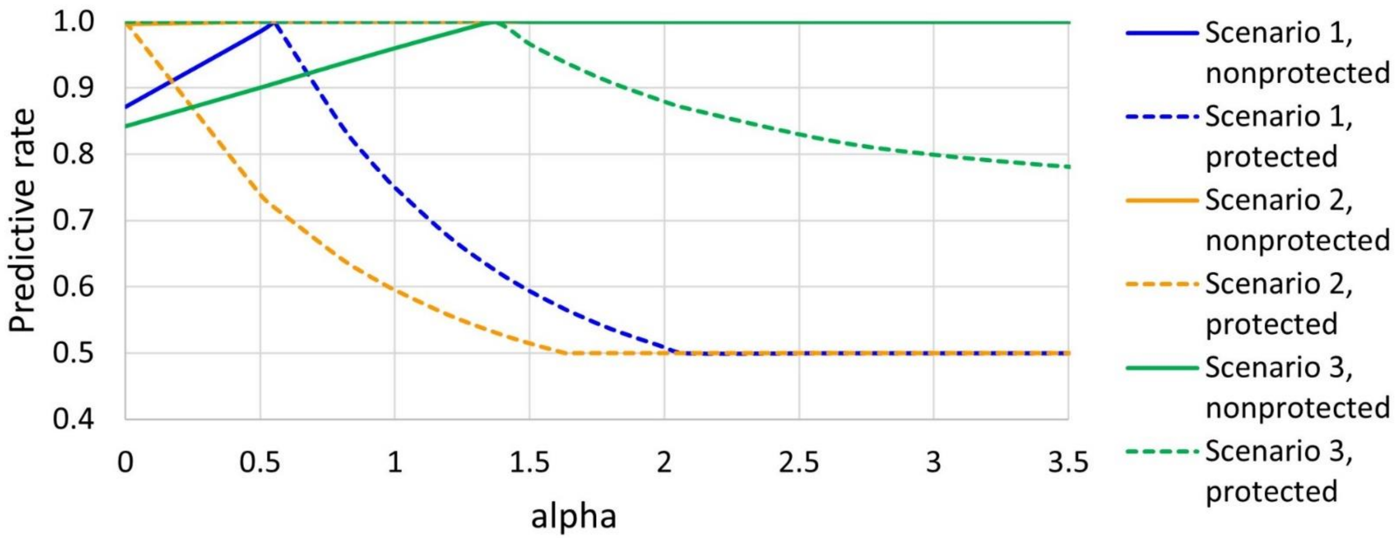
- Overall selection rate = **0.6** = overall qualification rate



- Only an **accuracy maximizing** solution (odds ratio = 1) yields equalized odds. Fairness not a factor.
- Nearly all odds ratios = 1 when selecting **more** individuals than are qualified.

Predictive Rate Parity

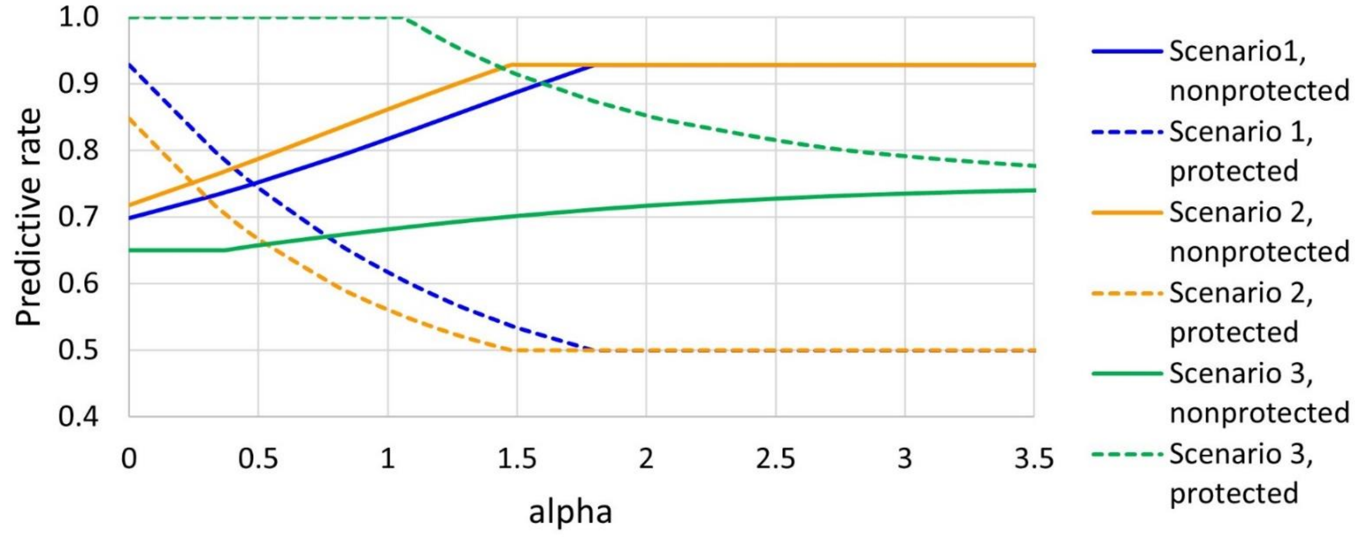
- Overall selection rate = **0.6** = overall qualification rate



- Higher predictive rates = **smaller** selection rates for protected group.
- Only an **accuracy maximizing** solution (pred rate = 1) yields predictive rate parity. Fairness not a factor.

Predictive Rate Parity

- Overall selection rate = **0.8** > overall qualification rate



- Nearly all predictive rates = 1 when selecting **fewer** individuals than are qualified.
- Predictive rate parity is a meaningful parity measure only when selecting **more** individuals than are qualified.

Conclusions

- Accounting for welfare
  - Alpha fairness (for suitable α) can result in **any** of the 3 types of parity, but usually when $\alpha < 1$.
  - So, parity is generally **less fair than proportional fairness**.

Conclusions

- Accounting for welfare
  - Alpha fairness (for suitable α) can result in **any** of the 3 types of parity, but usually when $\alpha < 1$.
  - So, parity is generally **less fair than proportional fairness**.
- Assessing parity metrics
  - Implications of alpha fairness depend heavily on **how many individuals are selected** relative to number qualified.
  - **Equalized odds** is a meaningful fairness measure only when selecting **fewer** individuals than are qualified.
  - **Equalized odds** is **less fair** (measured by α) than **demographic parity**.
  - **Predictive rate parity** is meaningful only when selecting **more** individuals than are qualified, which may be **unrealistic.**
    - Predictive rate parity may be relevant in the **parole** case (where lower recidivism corresponds to higher predictive rate) if one is willing to parole **unqualified** individuals.


Questions or
comments?